SUMMARY OF COMMENTS RECEIVED AND EMF STAFF RESPONSES

The comments we received ranged from the highly complementary to strongly critical. The EMF staff are gratified by the support expressed by eminent figures in the fields of public health, EMF research, epidemiology and risk evaluation:

"I highly appreciate the logical and disciplined way, in which you have analyzed the issues involved in evaluating the practical significance of a wealth of scientific work, which in the past had been undervalued. As your work shows, the reason for this is because earlier evaluations had been on the basis of an inappropriate set of criteria."
Theodor Abelin, MD MPH, Emeritus Professor, University of Bern

"This is by far the most comprehensive and ambitious risk evaluation on this topic that has been conducted. Not only is it comprehensive but also intelligently done. I also find that the process has been very carefully thought out with possibilities for all parties to participate and with a great openess. I realize that there has been comments on various aspects of the process, such as the choice of three assessors, but I am convinced both that you choose your methods very carefully and that whatever alternative you would have chosen had also generated comments." **Anders Ahlbom M.D. Ph.D., Karolinska Institute, Stockholm.**

"I would like to first congratulate you for preparing an excellent evaluation document..... The Evaluation is an excellent review by three experts. They have drawn conclusions based on their review and their background knowledge. They have made their determinations very transparent and understandable. They seemed reasonable and well considered."
Henry Anderson, M.D., Chief Medical Officer, Wisconsin Division of Public Health

"It is a very substantial document and clearly exhibits the great deal of thought and careful work that has gone into its preparation. In a number of respects it breaks new ground in the arena of public documents intended to inform the public and policy-makers about the hazards, if any, of EMFs.... [I]t is a valuable and unique contribution that the document 1) identifies differences in different scientific disciplinary approaches to the issues and considers with care cross-disciplinary implications; 2) considers that the users of the evaluation may well bring different ethical perspectives to the issues and care most about different aspects of the evaluation; 3) creates separate descriptions for pro and con interpretations of evidence to better illuminate both the evidence and the considerations that go into its evaluation; and 4) provides a variety of windows through which the reader can make his or her own interpretation of uncertainty in the judgements involved including both quantitative and qualitative descriptions by the evaluators, and the opportunity to observe the differences between the judgements of the individual evaluators and their explanations of how they formed their judgements."
Professors Robert Goble and Dale Hattis, Clark University, Worcester, MA.

"...a well written and well thought report that contributes greatly to the assessment of electric and magnetic fields as well as to the practice of policy development in environmental health. I think the Bayesian approach that you have developed is very consistent with how policy makers think about risk. More specifically your approach provides a great deal of transparency, which should allow all who are concerned about this issue to completely understand the assumptions and the logic that is built into your risk evaluation framework."
Lynn R. Goldman, MD MPH, Professor of Public Health, Johns Hopkins University.

"Overall, the review reflects a great deal of thoughtful work and, particularly in its presentation of ranges of uncertainties, presents a rare and more balanced view of the complicated nature of the risks being presented."
Ben Greenebaum Ph.D., Professor Emeritus/Adjunct Professor of Physics, University of Wisconsin.

"May I congratulate you and your colleagues on such a thoroughly professional and well thought out Evaluation. There is nothing like it in the UK or Europe and it should be a model for the authorities over here in their assessment of the EMF health effects."
Professor Denis L Henshaw, Professor of Physics, HH Will Physics Laboratory, Bristol, UK

"I found the report complete and thorough in its review of the major studies on EMF and disease endpoints... [A] splendid, even Herculean job"
Marc Lappe` Ph.D. Center for Ethics and Toxics

The document presents in great detail the process adopted for the evaluation and has the great merit of presenting in analytical detail the elements in favor and against any conclusion. This is a great merit as the process from scientific evidence to public policy is too often 'unintelligible'
C Magnani MD Ph.D. Professor University of Turin

"The evaluation is innovative, cogent, evenhanded, comprehensive and scholarly. Earlier efforts to make sense of complex data sets in environmental health began with narrative reviews, and later progressed to meta-analyses. This evaluation is a large step forward. Recognizing that everyone brings a prior set of beliefs to the assessment of a problem, it explicates the beliefs of the three evaluators, roughly quantitates them, and employs Bayes theorem to update their priors. The evaluators were trained in probability estimation, another innovation, and then confronted with the data. This is a huge improvement over the conventional narrative reviews, in which the reader must infer the positions of the reviewers.

The transforming of private hunches into public opinions is unquestionably progress. It permits the priors and the impact of newer data on these judgments to be examined and in turn evaluated."
Herbert L Needleman, MD. Professor of Psychiatry and Pediatrics, University of Pittsburg.

"This review of the scientific evidence on the health effects of low frequency EMF is the Cadillac of reviews, not just of EMF but of scientific reviews in general. It exceeds in thoroughness, thoughtfulness and sophistication any other review by a wide margin. Indeed, there are no comparable efforts which I am aware of on any subject.

The use of the Bayesian framework is especially appropriate for reviews of this nature. Most scientists, consciously or unconsciously, use Bayesian approaches to evaluating evidence, and the systematic recognition of this in the review under consideration enables the most coherence and compatibility with how scientific judgments are actually made. Moreover the care with which this was undertaken adds immeasurably to the product, making it not only a demonstration of principle but especially pertinent and useful.

I strongly endorse the scientific methods and reasoning used here, independently of any conclusions. It is outstanding."
David Ozonoff, MD, MPH. Chair of Environmental Health, Boston University School of Public Health.

"I find this document very well done, very instructive and very comprehensive. I am most impressed by the meticulous care given to the literature assessment and the major efforts invested to adhere to objectivity in interpreting this literature. It is very revealing to see each reviewer's judgment in the light of the evidence. It gives sound justification to their assessment.

On the negative side, I have little to say. In general, I found the "Statements to the Public" too cautious. Quite often, it is as if you are apologetic for having come to the conclusion that the review points in the direction of the existence of a risk."
Gilles Theriault, Professor, McGill University, Canada

". . .I wish to congratulate you. . . for the tremendous amount of work you did and your very remakable achievements. . . IARC has been sometimes too shy or exaggeratedly prudent in assessing risks based on evidence from some epidemiological studies or case-reports, with the result of minimizing them. This type of prudence does not encourage the adoption of measures of primary prevention and is not in favor of public health."
Lorenzo Tomatis, Formerly of IARC

Nevertheless, a sizable number of commenters criticized our draft with different degrees of disagreement with our methodology and our judgement for example:

""First the EMF *draft* Report on EMF overestimates the potential EMF risks. Second, a CDHS report that incorrectly assesses public health risks can have adverse consequences. Third it is important the CDHS correct the report…"
Abdelmonem A Affifi Ph.D Professor of Biostatistics and former Dean UCLA School of Public Health (as consultant for California utilities)

"1. The process uesed to prepare the Draft is neither an appropriate nor a reliable way to assess public health risks;

2. The Draft is neither consistent with the available science nor in agreement with other international reports prepared by independent experts;

3) The authors do not have the expertise in all of the relevant scientific disciplines to fully evaluate the EMF literature…

4) The risk communication messages have not been tested and are likely to be confusing and misleading, particularly to the general public
W. John Dawsey and other representatives of California Utilities

The logic of the of the analysis is very unclear and probably incorrect…..The panel pf the three reviewers is very poorly constituted...
Kenneth Foster Ph.D. Professor of Bioengineering and Electrical Engineering University of Pennsylvania

" I have examined in detail the reasoning and data used to reach this judgment as presented on pages 1-109 of Draft 3. I have concluded that the reasoning suffers from serious errors of logic and statistics, that there is repeated counting of the same data and subjective priors in the formation of conclusions, and that certain key data are not accounted for properly….These problems have led to overstatements of confidence about EMF effect in light of available data, and render much of Draft 3 invalid."
Sander Greenland M.A. M.S. Dr.P.H., C Stat. Professor of Epidemiology and Statistics UCLA (in consultation to California Utlities)

"Your draft report is in opposition to the results of the National Academy of Sciences, the Amercian Physical Society and the American Medical Association. Your tabular results of 50% causality for adult leukemia and 70% causality for adult brain cancer are beyond the pale!"
David Hafemeister Professor Emeritus Cal Poly San Luis Obispo and Study Director, American Physical Society Study "Power line Fields and Public Health"

The weight of evidence from rodent bioassays was minimized….The CDHS reviewers chose to put little weight on the absence of a biophysical mechanism….The CDHS reviewers conducted a pseudo-Bayesian analysis, which did not permit an objective(or quantitative) of how bias, confounding or measurement error might play into the epidemiologic results…. The CDHS review suffered from a lack of diversity in its scientific panel.
Robert Kavet Sc.D. Electic Power Research Institute

"The confidence that scientists and publich health officials can place in the EMF Program judgements is seriously undermined by the problems in the clarity of these reports and their conclusions, in the condut of the risk evaluation and in the failure of the EMF Program to advance public policy appropriate to the weight of the scientific evidence. Therefore, the evaluation and methods of the draft reports should be revised and redone by a review team wholly independent of the EMF program.
Richard M. Loughery, Director, Environmental Activities and EMF Issue Manager Edison Electric Institute.

"The draft report is overly complex, confusing, misleading and difficult to understand. Unless corrected, the weakness in the risk communication approach will make it more difficult to establish appropriate policies.."
Jack Sahl Ph.D. J Sahl & Associates, Peter Valberg Ph.D. Gradient Corp. (consulting for California utilities)

"The study has clearly involved an extraordinary amount of work on a difficult subject which has never before had such attention. But I believe that it is flawed(although perhaps not beyond redemption)"
Richard Wilson Ph.D Professor Harvard University

Among the more than seventy comments received, the following are the major objections we identified:

1. This whole exercise is a waste of taxpayers' money

2. We do not understand the reasons for this Bayesian approach or we disagree on the way you did it.

3. Your reviewers' panel is too small, unrepresentative and not sufficiently independent.

4. Your conclusions are out of line with the authoritative IARC evaluation

5. The prior is not a single value, but a distribution of values and they depend on the definition of "exposed" and "non-exposed"

6. We don't agree with your use of Hills's criteria of causality

7. Your evaluation is largely drive by what you call "consistency" of the epidemiological results. This attribute is already captured by the point estimate and variance of the relative risk as assessed by a pooled or meta-analysis.

8. You rely heavily on an inefficient and misleading statistical tool, the sign test.

9. you double count the evidence when you consider consistency and homogeneity of the epidemiological results

10. The belief that the sign test takes into account the effects of random biases is ridiculous.

11. You overplay the significance of your own results on spontanepous abortion.

12. You underplay the importance of possible selection and non-participation bias.

13. Twenty years of research should have brought more definite results, if there was a true underlying risk.

14. It is unscientific to use a “degree of confidence” in the existence of risk. Either exposure entails a risk or it doesn’t.

15. You relied too much on epidemiology and dismissed too easily the other streams of evidence

16. This whole debate ignores the basic laws of electromagnetism.

17. Power frequency photons do not have enough energy to constitute a cancer risk

18. We have problems with your statement to the General Public.

19. Your conclusions are different from what I see in the same body of evidence

Regarding the last point, we acknowledge that, when evaluating a large, complex and not entirely consistent body of evidence, different objective and reasonable reviewers may reach different conclusion. We are aware that the EMF debate will continue after this evaluation and that decision makers will hear diverse opinions from many quarters. We do not ignore the fact that many commenters (Afifi, Dawsey, Wilson) have indicated that they interpret the evidence differently and we respect their judgement, but we do not feel compelled to change our evaluation to reflect other people's conclusions.

On the other hand, other commenters have disagreed with our motivations, approach and use of logical and scientific arguments. We have reviewed all these comment thoroughly, aware of the fact that if these criticisms turned out to be well founded, our evaluation would be severely undermined. For this reason, we dedicated most of our responses to this type of objections. Since the same observation may have been made by more than one person, we have grouped our responses by topic.

1. **EMF is a dead issue. There have been many evaluations concluding that there is no conclusive evidence of risk. This evaluation is unwarranted, why then was it done?**

   The simplest answer to this question is that the California Public Utilities Commission asked the Department of Health Services in 1993 to address this issue and the stakeholders who oversaw the project asked the Department to do a risk evaluation. However judging by the tone of the reviewers who raised this issue they obviously disagree with the assignment given by the CPUC because they concur with the 1995 statement of the American Physical Society, and are apparently virtually certain that EMFs are harmless and that any effort directed to them is a waste of time and money. We point out that most recently, the National Institutes of Environmental Health Sciences(1998), the British Radiation Board (2001) and the International Agency for Research in Cancer (2001) also expended public moneys on risk evaluations and all concluded that EMFs were possible carcinogens on the basis of the childhood leukemia evidence. Indeed, the World Health Organization is now recommending "prudent avoidance" So agencies whose role it is to provide societal guidance on health matters, are like us, in disagreement with the American Physical Society, which is a professional organization whose main mission and expertise is not related to health. Our risk evaluation was packaged to be useful in policy analyses that asked the question: "how confident must one be of how much of an effect before inexpensive and expensive EMF avoidance measures should be taken." The other risk evaluations are not packaged to be helpful in answering this question. Our risk evaluation is not redundant in format to the other risk evaluations that began after our initial assignment was given. This is why we used the rate payers money entrusted to us by the CPUC to do yet another risk evaluation.

2. **How Does Your "Qualitative Bayes" Evaluation Compare to Traditional Risk Evaluations and To a Quantitative Bayes Model. Does the "Qualitative Bayes" Evaluation always produce larger degrees of confidence than a traditional evaluation would? Why don't you do a quantitative Bayes approach?**

   In brief, we wanted a more structured evaluation than those conducted under established guidelines (where evidence is briefly discussed among a panel of experts, who then vote on a conclusion) and we saw merits in the Bayesian approach, that considers the a priori likelihood of a hypothesis and the likelihood of evidence existing under both the hypothesis being considered and the null hypothesis. However, were we convinced that, given the paucity of the quantitative evidence, a formal, quantitative Bayesian analysis would have been based on so many subjective assumption as to make a quantitative answer meaningless. For a detailed discussion, please see Appendix 1.

3. **Other reviews were conducted by large panels comprised by representatives of many disciplines, employed by different agencies. How can only three DHS epidemiological reviewers to provide a knowledgeable and representative judgement about the potential hazards of EMFs? (Names.....)**

   This comment raised several issues, First, is a sample of three sufficient to provide a "representative" judgement. Second are these particular three reviewers representative or aberrant in their views. Third is the expertise possessed by these particular reviewers wide enough to properly judge the many streams of evidence?

   It is not at all clear who should be included in the universe that one should be trying to represent. Should it be scientists who have read and understood all the articles in all the streams of evidence? We suspect that none of the members of expert committees assembled to date would meet this criterion. Should it be scientists who have published several articles in one or the other fields? Should it be scientists who have tried to systematically review the several streams of literature? Should it be scientists who have not taken hardened stands in the controversy?

   We have been following this field for more than a decade and it is our impression that scientific judgements in this field would not be graphed in a smooth bell shaped curve with one central tendency for degree of confidence. There are various camps. The January 2001 special issue of Bioelectromagnetics, covers the special epidemiological workshop that we sponsored that involved published researchers in the field selected to represent the full range of opinions. In it we document the degrees of confidence among attendees, that the childhood brain cancer epidemiology revealed a causal relationship. Two groups emerged, one that tended to doubt a causal relationship and one which tended to give causality credence (the three DHS reviewers tend to side with the doubters on this endpoint). There is no representative sample, of any size, that would yield one degree of confidence that could properly represent the judgements of divided knowledgeable scientists in the EMF arena.

If competent researchers have a range of opinion on this matter, we believe that decision-makers need to know that range and relative frequency of opinion not the weighted average between two or more different camps. We plan to modify the risk evaluation to provide greater space to document evidence for the range of judgements in this field for the PUC and to make clear where the DHS reviewers stand relative to other expert committees. We think that our reviewers fall in the upper 25th percentile as to degrees of confidence. Time will tell whether this minority opinion turns out to be the right one.

The California Public Utilities Commission and the overseeing stakeholders asked the Department of Health Services to provide a judgement about a controversial area. The Department assigned its most knowledgeable scientists to make this judgement, much as a district court will assign a judge to evaluate the appeal of a controversial court case. The focus is on responsibility and competence not on representativeness of what the universe of judges might have decided. Since the bulk of the controversy relates to the epidemiological evidence and the biophysical "impossibility" arguments it seems appropriate that the three reviewers included scientists with backgrounds in physics and epidemiology. The fact that they did not let the largely null animal pathology pull their confidence down as much as some commentors would like, does not derive from a lack of understanding of animal bioassays, and the fact that mechanistic experimental results did not pull up their degree of confidence does not derive from a lack of understanding of the mechanistic studies. Indeed the team contracted with experts in animal studies and biophysics and mechanistic studies to review all the new studies that had emerged since the deadline of the NIEHS review to make sure that this expertise in summarizing the literature was available to influence their judgements. For each epidemiological chapter, and the animal and mechanistic chapters other DHS epidemiologists or toxicologists reviewed the same body of research papers and commented on the core reviewers write-up. Some fifteen additional DHS scientists were involved in this effort to revise the internal first draft. Two outside senior epidemiologists were asked to critique an earlier version of the qualitative Bayes approach and major changes were made as a result of their comments. Additionally the second draft of the evaluation was sent to the multi-disciplinary Science Advisory Panel and changes were made that are reflected in the third draft.

There are pros and cons when comparing a DHS staff evaluation document to the external committee documents prepared by the NIEHS, and IARC. Those committees are larger, but their conclusions will be influenced by the relative number of believers or doubters selected by the staff to attend the process and by the formats used for counting the votes and forcing a consensus. The reports were written in the course of a week, under great time pressure, with different chapters prepared by different disciplinary groups.

The DHS staff document has the value of extensive comment and revision within the working group and outside reviewers. The fact that reviewer 2 was the bureaucratic superior of reviewers 1 and 3 does not seem to have induced unanimity, since each reviewer was encouraged to use their best scientific judgement. The fact that all the reviewers have survived in an institution whose mandate is to protect the public's health is likely to have selected for approaches to evidence that would be different from scientists who have survived in other kinds of environments (reference to south Carolina thesis on this). Ostensibly, however, this is why the CPUC turned to CDHS to do this evaluation in the first place so that this issue of what kind of scientists to select would become moot.

In summary, we conclude that the three DHS reviewers particularly with the outside help of which they availed themselves, are professionally qualified to have made this evaluation and that selecting additional scientists from inside or outside of the Department to review the extensive literature in the format we have used is neither logistically nor economically feasible. As stated above, we plan to deal with the "representativeness" issue by devoting more space to the opinions of other reviews and how we stand in relation to them.

4. Even when you used the IARC guidelines, your evaluation is often substantially different from that of IARC and NIEHS. How do you explain this? (Names .....)

We believe that part of the difference is due the position of the DHS reviewers and part is due to the process followed and the assumption of a 'default position'. Please see Appendix 2 for a detailed discussion.

5. **One should not have *one* prior, as you reviewers did, but a distribution of prior probabilities for relative risks at each level of exposure. This problem is also reflected in the decision models. Anchoring the prior probability in the first and ninety fifth percentile of the United States population raises a number of problems: These bench marks have no biological linkage to laboratory experiments, they are not California values, they do not relate to European values where the exposure distribution is different. (Greenland).**

Our experienced decision analysts both advised us to simplify and approximate the posterior dose response by formulating a prior probability for a reasonable range relative risks and use this as a weight for exploring the consequences of assuming several possible dose response types of various steepness. If there was more information and understanding of dose response the approach proposed by Professor Greenland would be more compelling and would allow him to predict effects in any location for which he knew the distribution of "exposure".

But in fact we do not have that kind of understanding, and it is not at all clear that pursuing the more elaborate formalization would be a closer approximation to the truth, given all the uncertainties. It was clear to our consultants that it would make the exploration of options much less tractable and understandable.

We anchored our priors to the first and ninety fifth percentile of the US population to operationalize a more general concept. The issue here is the prior credibility that the usual ambient range of an agent could produce an effect that even an epidemiologist could see. Anchoring on the 95th and 1st percentiles, whatever these may be is one way of operationalizing this idea. It would work for the very few other agents that produce epidemiologically observable effects in the general environment even if the true effect only began at the 96th or 97th percentile.. Once anchored in the US it could be translated to the UK (say the 10th and the 99.9th percentile). We referred to the US distribution of personal exposure because a published study existed. The only California data for personal exposure are those on 600 and 1000 pregnant women in the studies respectively of Lee et al and Li et al whose smoothed distribution have not been published, but are similar to the US distribution.

As for anchoring in experimental data, it would be the rare environmental agent, that demonstrated positive experimental results in a small number of laboratory animals at ambient exposure levels. Like the topic of particulate air pollution and mortality, we have persistent epidemiological associations of modest size with little or no animal pathology experiments to provide an anchor.

Decision makers are at a decision juncture now because of the NIEHS, IARC and California summary of the first long phase of research in this area. They can wait for more information, or on the basis of current information they can decide to do nothing, do no and low cost avoidance or expensive avoidance. By insisting that no policy analysis is possible until information is available to do an academically "correct" prior probability analysis. Professor Greenland is in effect arguing for the "wait for more information" risk management option because without the additional information we are completely at sea and can make no policy analysis. We disagree.

6. **The Hill's criteria are not specific rules to establish causality, as pointed out by Bradford Hill himself. In a recent textbook Greenland and Rothman have criticized them as outdated. Yet you used them mechanically and unwisely organized your thoughts around them. (Greenland)**

In our guidelines we referenced the Rothman and Greenland critique of Hills criteria and received critical comment about that from some reviewers of the guidelines. As we did the actual risk evaluation we began to realize, that with refinement of definitions, Hills criteria served us well. We let "Plausibility" stand for mechanistic research and "Experimental" stand for animal pathology research, two distinct streams of evidence. "Analogy" was always a possibility but never came up and its absence was given no weight. "Specificity" invoked two different concepts for us. One was whether one had an a priori belief that an agent associated with one specific subtype of cancer (say) was more believable as a cause than an agent that was associated with several types of cancer. In the early days this kind of "specificity" was supposed to boost the credibility of the causal hypothesis. More recently we have recognized that most carcinogens cause more than one type of cancer and many people now believe that this property carries little weight one way or the other. Nonetheless we have heard arguments that EMFs association with various cancers is not specific, and for some people this lowers their belief in causality. Reviewer 1, but not Reviewers 2 and 3, thought that, since, unlike other agents, EMFs reach all organs, there was *a priori* reason to believe that they would cause a bigger variety of diseases than the typical chemical agent,

which is normally constrained to a specific biochemical path. Hence we retained this category. The other concept was, that once one has a posteriori demonstrated that an agent has one kind of health effect, one may be more willing to believe that it could increase the risk of another diseases. All three reviewers thought so. So we added another category "associations with other diseases". This seemed like double counting to Professor Greenland. While we think the concepts are different we will collapse these categories into one called "Specificity and Associations with other Diseases".

7. **Your evaluation is largely driven by an attribute you call "CONSISTENCY". This attribute is redundant since it is already captured by the variance and effect measure of a pooled or meta-analytic summary of the studies (Greenland).**

   We disagree. By looking at the proportion of studies that reported an increase in risk, we are asking if there have been replications of the studies in different setting and with different methodologies most of which tend to show a result in the same direction.. It seems that Greenland's degree of confidence would be indifferent to results that came from one enormous study or from a whole series of moderately sized studies whose aggregate number of cases and non cases equaled that found in the enormous study. From a purely statistical point of view this might be equivalent, but it ignores the value of replication.

8. **To measure "consistency", you use the sign test. This is "inefficient and potentially misleading" (Greenland) and a "statistical approach that contradicts standard epidemiologic and statistical practices for evaluating research. It does not reflect the inherent strengths and limitations of epidemiologic data" (Kavet).**

   We acknowledge that the sign test is a crude statistical tool, but we think is valuable because it allows to examine a whole body of epidemiological evidence, including studies of different design and studies for which the original crude data are not available. We do not believe it is misleading. For a detailed discussion of the pro and cons of the sign test, please see appendix 3.

9. **Beside "consistency", you consider an attribute that you call "homogeneity". These two attribute overlap and you are double counting the evidence (Greenland).**

   We called "consistency" the proportion of studies falling above an OR of 1.0 and "homogeneity" how tightly the studies falling above an OR of 1.0 clustered around each other. We realize that if you don't have consistency you cannot have homogeneity, but you can have consistency and not homogeneity so they are different ideas. We will clarify this in the discussion in chapter 7.

10. **Reviewer 1 states that unidentified biases "are random events that are accounted for by an appropriate statistical test, such as determining the p-value using a sign test." If it were true, observational researchers would never have to worry about biases, knowing that the statistics took them into account.**

    Reviewer 1 responds: the statement (which was taken out of context) refers only to unidentified biases which may affect the risk estimate either way. To explain this, consider the following analogy: I have a large number of playing dice and I have to determine whether they are suitable for gambling in a licensed establishment. I examine them and I am not able to distinguish or even suspect of any irregularities (identified biases). I throw one die a large number of times and I observe an extreme number throws returning an even number. I conclude that that die is biased. I am now asked whether the dice are biased because they are of poor quality (unidentified, random biases) or because someone deliberately tampered with them (causality). I reason that, if someone had tampered with them in order to give himself a winning margin, he would have done so in a systematic manner, to increase the predictability of a given result. Therefore, I throw each die once and only once. If they are biased because they are cheaply made and unevenly worn out, I expect to see an approximately equal number of odds and even numbers. If an unusually high number of them land on an even number, I have reason to believe that they have been tampered with. And how do I determine whether the number of even numbers is "unusually high"? With a sign test.

    However, there is one case in which the sign test is inadequate: if a number of the die are biased because of poor manufacturing AND were made by the same inaccurate machine (***shared*** random bias). In this case, all the dice made by that machine are biased in the same direction, although

unintentionally. A sign test, detecting an extreme pattern, may inaccurately suggest causality.

In the EMF case, many studies were conducted in the same nation (USA) and several other in a group of culturally and socially similar nations (the Scandinavian countries). If an unidentified bias affected all US studies or all the Scandinavian studies, the effect of this bias on the pattern of result could be considerable. Still, we have no clue as to whether this shared bias would affect the result upward or downward. Prudence may suggest to adopt the "worst case" scenario, but the question is "which way is 'worse'?". An academic researcher may be more inclined not to report a positive finding if it is still uncertain. A public health official may deem it more prudent to signal a possible risk, even if it may turn out to be a false alarm.

Nevertheless, I have considered a scenario in which all the measurement-based US studies are regarded as one trial with a negative outcome (i.e., excluding the registry based Wertheimer and Leeper study). Then, we would have 13 studies with two successes (i.e., RR< 1). This is still a very extreme pattern (P<0. 0112). We can push this further now by including the Canadian studies with the US studies. Now we have 11 trials and 2 successes. The p value is larger (p<0.0327), but still suggesting that causality is the most likely explanation. To consider an even more extreme scenario, we regard the Scandinavian studies conducted in the 1990s as one trial (the 1986 Tomenius study is methodologically so different that it is not comparable to the others, even if it was conducted in a Scandinavian country). We also assume that all these results are entirely driven by bias. We now have 9 trials and 3 successes (p<0.2539). Our confidence that this pattern is not due to chance is now only 75%. This figure is still within the margin of error I have placed around my estimate. Given the extreme assumption behind this figure, I remain of the opinion that causality remains the most reasonable explanation.

Yet another way to consider a "worst case" scenario is to assume that, even if there is no true association, the probability of any study returning a risk estimate > 1 is 75%, instead of 50%, because of unidentified biases affecting the risk estimate upward. Even under this assumption, the probability that out of 15 studies no more than one returned a risk estimate ≤ 1 is only 0.0802 (to be evenhanded, we should also consider the opposite scenario, in which biases push the risk estimate downward and the probability of a result >1 is only 25%, in which case the p-value is < 0.0000).

Reviewer 2 Responds: (I will do this later this week)

11. Are you giving to much weight to two unpublished miscarriage studies associated with your program?

The state legislature funded a large prospective study of electric blanket use with a nested case control study of the association of personal 24 hour EMF exposure and miscarriage. Preliminary results of this study in the early 1990s led the EMF program and a specially convened group of experts to recommend that the study be replicated prospectively and the EMF program funded a million dollar 1000 woman study in which personal 24 hour EMF exposure was to be ascertained prospectively before miscarriage. At the time of the risk evaluation, these studies received a special external peer review so we felt comfortable considering their results. Subsequently both articles with their original conclusions, but somewhat shortened, have received additional peer review and will appear in the January 2002 issue of the journal Epidemiology. We believe, knowing these results, that it would have been unwise to ignore them just because they were not yet published. Studies of the earlier high exposure VDTs a study in Finland and conflicting studies of electric blankets had suggested before that something might be going on. Objectively speaking, the Lee and Li studies are elaborate state of the art studies and aside from the fact that two of the three reviewers were directly involved with the Lee et al. study and all three reviewers provide advice on the Li et al. study, the weight we put on their results relates to their size and quality not to our personal involvement with them. In fact, Reviewer 1's opinion is not based by the main findings of Li 2000 and Lee 2001 (see Draft 3, pg. 206, lines 26-28).

The suggestion by Kavet that Reviewer 1 has expressed an opinion inconsistent with his previously published article is incorrect. The article expressed the opinion the VDT work was probably not a hazard, because of the very low exposure from the modern terminals, but that exposure from old model VDTs and other ELF sources could not be declared safe.

12. You underplayed the possibility that selection and non-participation bias could be responsible for the observed associations.

Those who doubt the epidemiological associations between residential EMF and various cancers believe that their is a possibility that selection and non-participation bias may have affected the results of many epidemiological studies to the extent these biases could, in combination with chance and

confounding, explain away the association between EMFs and cancer. The main basis for this claim is an article by Hatch et al, which was unpublished at the time of our evaluation. These authors found that the odds ratio for childhood leukemia among those living in homes with very high current configurations increased by 23% when 107 "partial participants" were excluded. "Partial participants" tended to be characterized by lower socioeconomic status than subjects who participated fully, suggesting possible selection bias. Hatch's results are in contrast to Savitz's (1988), who, comparing the complete wire code data set and the measured field data (limited by some non-participation) estimated that, had the participation been higher, the relative risk would have been stronger. Both Hatch's and Savitz's conclusions are speculative, but they point out how it would be a mistake to assume that observations relating to one study apply to the whole body of evidence.

Jaffa suggests that, in the Swedish childhood leukemia study, the discrepancy between the results obtained using contemporaneous magnetic fields measurements (which show no association) and those obtained from historical magnetic field calculation (which show an association) cannot be explained simply by misclassification of exposure and that a more likely explanation is that that historical calculations introduced some sort of bias. We disagree, but we note that the argument is largely moot. Both the Wartenberg meta-analysis and the Ahlbom pooled analysis have carried out sensitivity analyses which indicate that if the Swedish study results are excluded, the association between magnetic field and childhood leukemia is not significantly weakened.

Ahlbom et al (2001) evaluated the effect of possible selection and non-participation bias on the findings of their pooled analysis of EMFs and childhood leukemia. They noted that excluding groups of studies from the same geographical region (which might be suspected to share similar associations between demographic characteristic and environmental exposures, including EMF), did not significantly change the results of the analysis. In particular, the Nordic studies which were cohort studies or nested case control studies that did not require the cooperation of subjects and were therefore not prone to selection bias, showed similar results to those from random digit dial case control studies which are potentially prone to selection bias:

	OR for exposure $ 0.4 μT
All studies	2.0 (1.3-3.1)
US excluded	1.7 (1.0-2.8)
Canada excl.	2.1 (1.3-3.6)
Nordic countries	2.1 (0.9-4.9)
Rest of the world	1.9 (1.1-3.2)

Ahlbom et al concluded that selection and non-participation bias:

May be of importance in some studies

Is of magnitude and nature not different from other research areas

Is, alone, unlikely to explain EMF association

We note that, if we assume (against the evidence) that all the study results were inflated by selection or non participation bias, this would not change the results of our sign test, which only considers the existence of an association, not its size. We should also point out with regard to selection bias and social class that there are cohort studies of childhood leukemia, adult leukemia, adult brain cancer and Lou Gehrig's disease that show positive associations even though selection bias is not a problem of cohort studies. Therefore selection bias is not a credible explanation of those results.

Nonetheless critics point that it is a common occurrence in epidemiological studies to have a higher participation rate on the part of better educated subject and, when contact is made by telephone, a lower selection rate of low income subjects who may not have a telephone or are away from home working several jobs or suspicious of institutions that they feel have failed them.. Assuming that higher SES subjects are less likely to live near powerlines, these critics conclude that the controls have lower exposure because they are not representative of the population at large, and not because lower exposure is associated with lower disease risk. The logic of this criticism involves the following assumptions:

Low income is differentially under-represented in controls as compared to cases and

Homes closer to power lines are often occupied by lower income families and/or

Lower income people also have more exposure to other sources of EMFs.

If these assumptions were correct, selection bias based on income will produce a falsely low prevalence of VHCC homes and a falsely low distribution of exposure in the control series and distort a truly null association so that EMFs falsely appear to convey an increased risk of whatever disease is being studied.

Greenland quotes an article by Bracken et al (Am J Epidemiol 1998; 148; 467-74), which reports a strong correlation between some SES indicators (woman's occupation, house value) and the very high current configuration (VHCC) wire code configuration. It was also pointed out, that wire code persists as a predictor of childhood leukemia in Greenland's pooled analysis even when controlling for measured magnetic fields. This suggested that wire code has an independent effect, or that this was proof of a selection bias related to wire code.

To address these concerns, we have explored two very large data sets collected in the San Francisco Bay Area. We found some suggestive, but inconclusive evidence that participation is correlated to wire code status. However this correlation was not due to SES as suggested by Bracken (1998) and Hatch (2001). Furthermore, we found no evidence of an association between family income and measured EMF exposure. Analysis of these data is presented in detail in appendix 4.

Neither these, nor the Bracken data quoted by Greenland pertain to children in the areas where the childhood leukemia studies were conducted (they all refer to women of child bearing age), but they strongly argue against the suggestion that the association between low SES and higher exposure is strong and common. Thus even if study participation varied by income, this would not distort the distribution by personal magnetic field exposure in the study subjects. On the basis of the published literature and the as yet unpublished information from the two Bay Area pregnancy studies we see little objective evidence that selection bias explains the epidemiological evidence linking measured magnetic fields and disease in case control studies, much less in cohort studies.

13. Doesn't the 20 years and many millions of dollars spent on EMF research without a definitive result suggest to you that there is no real effect to be found ? (Sahl and Dawsey)

This does not pull down our confidence very much. Funding for EMF research has come in fits and starts as public attention has lead to time limited research efforts such as our own, the New York State Program and the recent federal program. The Department of Energy program, which was a small relatively steady program, has come to an end and the Electric Power Research Institute program is much reduced. Many of the research programs were not policy oriented or targeted, so that there are few researchers who have been able to consistently follow a policy relevant line of research through a series of research iterations. Thus the length of time and the amount of money spent is somewhat deceptive. The "war on cancer" research program started by President Nixon has been going rather steadily for more than 30 years with modest benefits only recently accruing.

14. Why are you using the "unscientific" degree of confidence of causality to express your judgements? (Dawsey)

Although the word "unscientific" is sometimes used as a pejorative term, it literally means, "not using the methods of science". We have used accepted methods for presenting the evidence and evaluating it. There is, however, no known method that assures unanimity, for combining streams of scientific evidence for deriving a judgement as to hazard. While there are rules of evidence used by USEPA and IARC, they are more like legal rules of evidence and cannot be characterized as scientific. It is true that many scientists feel uncomfortable couching their judgements in probabilistic terms and tend to dichotomize agents into those for which they are virtually certain as to their hazard and those for which the evidence of hazard is not conclusive. This spares them the task of distinguishing between agents that they are 11% sure of from those that they are 89% sure of and this projects an aura of scientific definitiveness, which is not helpful for making decisions other than the decision to "take no action until you are certain". Because we did not want to preclude action in the face of uncertainty we funded decision analytic models and packaged our judgements in this probabilistic way.

15. Why so much emphasis on Epidemiology?

Several commenters have noted our heavy dependence on epidemiological evidence and some of them have questioned the validity of our conclusions

because of the scant (or, some say, non-existing) support we derive from mechanistic and animal evidence. We acknowledge that our evaluation is driven by the human (epidemiological) evidence. We are not alone in this. The IARC guidelines place human evidence above all others. If human evidence is sufficient, an agent can be classified as Group four (human carcinogen) even if no animal evidence exists. Note also that no other recent authoritative evaluation (NIEHS, NRPB, IARC) found animal or mechanistic evidence convincing, yet they did not dismiss the possibility of EMF health effect on the strength of these non-supporting results. The question of whether epidemiological evidence alone is convincing is of course, at the core of the debate.

16. How did you calculate the attributable fraction?

Prof. Greenland criticized on several technical grounds our method of constructing an approximate dose-response curve by stratifying the risk estimates on a number of discrete exposure bands. We accept that ours was not a rigorous approach, but we believe that, given the uncertainties attached to the data, it would have been provided a reasonably acceptable estimate. Instead of trying to calculate this directly we simply take Greenland's estimate, divide it by 3 to account for the weaker associations for some of the implicated diseases and demonstrate that even a 1% population attributable risk percent would generate a population burden that is larger than that calculated for many regulated chemicals. We leave to the powergrid and school policy analyses the estimates of possible decrements in case load from actual mitigations. The equation in draft 3....is therefore deleted and the questions raised by Dr Kavet become moot.

17. You show little regard for the science of Physics. NOT one epidemiologist has stood up and said, my result is so reliable and my interpretation so correct, that I am prepared to say that there is something new going on that is contradiction to the work of Maxwell, Einstein and Dirac. The letters EMF was known to all of them as "Electro Motive Force" and its use otherwise shows a lack of appreciation of the historical context. (Wilson)

We never argue that the only way for EMFs to affect biological processes is through some unknown mechanism that overrides the laws of electromagnetism. The argument is about whether a weak signal can be perceived above strong noise. We argue that it is possible, because the signal is qualitatively different from the noise, being coherent in time and space.

We know that the acronym EMF is used to indicate electromotive force (usually, in lower case letters). EMF also stands for Electronic Music Foundation and is the name of a rock band as well as an antique firearms importer and an Internet service provider. It is also used by respectable organizations including the WHO to denote electromagnetic fields. In this document it is clearly defined as an abbreviation for electric AND magnetic fields.

18. In order for any agent, radiation or chemical, to be a human or animal carcinogen the agent must have the ability to interact with cellular systems and macromolecules. Upon absorption there must be enough energy to disrupt chemical bonds and initiate a destabilizing effect on the genome. (Cleaver).

Strictly speaking this is true. In fact, we prefer to use the expression "risk factor for cancer" rather than carcinogen. To be a risk factor, an agent does not need to interact at the molecular level. For example, it is believed that a low fiber diet is a risk factor for bowel cancer for the simple reason that it allows waste products to remain in the bowel for longer periods than if large amounts of fiber were ingested. A depression in the immune system and hormonal imbalances are capable of incresing the risk of some cancers without acting on the genome.

RESPONSES TO STRUCTURED QUESTIONS

In sending out the draft for public review, the EMF Program had prepared a list of eight questions on were regarded as important aspects of this evaluation. The following table lists the questions, the commenters responses, when provided and the staff's comments and replies.

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
Name of respondent	**Response to question**	**Staff reply**
Theodor Abelin	I fully agree with your position that theories should not be used to discount evidence, in particular if it is well documented and consistently found. In the interplay of theory and practice, such as has been typical in physics, theory has to be consistently reformulated to take into account new evidence. Otherwise we do not deal with science, but with religious or esoteric belief systems.	We acknowledge your support for our point of view
Ross Adey	I agree. Objections from physicists arise in equilibrium thermodynamics involving a basic tenet that perturbations with energies below atomic thermal collision energies cannot be effective stimuli. There are many observations to the contrary in biological systems, as for example in the ear, where the auditory threshold involves a receptor displacement of 10^{-11} meters, or the diameter of a single hydrogen atom and substantially below the collision energies of receptor atoms and molecules. The doctrinaire attitude of these physicists is that, "Since your observations do not fit my models, therefore they are artifacts." Similar objections from biologists also typically arise in inappropriate models of threshold sensitivities, based in equilibrium thermodynamics and ignoring the overwhelming evidence for nonequilibrium, nonlinear organization in biosystems. This involves coherent energy states and highly cooperative transitions, with good and growing theoretical and experimental evidence for sensitivities below thermal thresholds, and very importantly, with evidence that tissue sensitivities are set by populations of elements and not by a single receptor.	We acknowledge your support for our point of view
Anders Ahlbom	I fully agree.	We acknowledge your support for our point of view
David Bates	I agree that arguments based on physics and simplified models should be disregarded. I was Vice-Chairman of a Science Advisory Board Committee to critique an EPA report on EMF; we heard evidence that there was clear	We acknowledge your support for our point of view

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
	evidence that weak 50 Hz fields could influence intact systems although several senior physicists told us that no tissue effects were possible. We were influenced from experiments on the effects of radar on bird migration-, on demonstrated effects in expediting bone healing; and on effects on fish. All of these demonstrated that the arguments from purer physics were not applicable.	
Carl Blackman	I agree that physical theories regarding the likelihood that EMFs at residential and occupational levels could cause bio effects should not be used to discount evidence. To do otherwise would be contrary to use of the scientific method.	We acknowledge your support for our point of view
Bowman	Questions 1, 3 and 4 all deal with the difficult aspect of balancing mechanistic evidence from theoretical and laboratory research results (generally lacking evidence of EMF bioeffects) with the epidemiological research results that provide evidence of EMF health effects. The low weight given to the absence of support from simplified biophysical models is reasonable.	We acknowledge your support for our point of view
Shan Cretin	I agree that the lack of biophysical models currently able to explain an effect should not create a prior so strong that epidemiological data could not possibly affect the posterior. However, I do think that marginal or equivocal epidemiological data will have stronger weight in the presence of a well-understood mechanism. Is this the same thing as saying that the lack of a theory leads me to discount evidence? I think so, in that the conditional probability of there being a real EMF affect given the same set of evidence AND a plausible mechanism is bigger than the conditional probability of there being an effect given the same set of evidence and NO plausible mechanism.	We acknowledge your support for our point of view and agree with your further comments. In setting the prior, we did no consider any evidence obtained from EMF-targeted studies. The biophysical models are based only on prior knowledge of physics and biology, therefore, even if developed during the EMF debate. Therefore, they can be used to evaluate the a priori plausibility of the hypothesis. According to this models, the hypothesis should have no credibility at all. We disagree, because we do not believe these models to be less than adequate. However, they do suggest that the energy of environmental fields is, if not necessarily too low to be perceived, certainly very low, and this has been considered by all three reviewers an argument in favor of a lower prior that would have been otherwise the case.
John Dawsey *et. al.*	We disagree with your use of biophysical evidence. First you misconstrue the argument, saying that since 'biophysics does not prove that EMF is safe, then this stream of evidence is not valuable'. You also discount biophysical theory by over-looking the substantial base of direct, reproducible experimental 'observation' that was used to construct these theories. The point isn't that your prior degrees of confidence are too low; it is that the three	In this question, we do not refer to the science of biophysics, but to the use of biophysical models to predict, a priori, that no effect is possible. We believe that these models have been proven (often by the same authors) to be susceptible of improvement and therefore not authoritiative enough to dismiss empirical evidence out of hand. We did consider biophysical plausibility in setting our priors. Specifically,

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
	reviewers give too much weight to a highly selected set of 'new epidemiological information'. You fail to recognize the added importance of biophysical plausibility when the epidemiology conclusions are based on small numbers and weak effects, and no specific magnetic field parameter has been identified. In sum, you have consistently underestimated the value and relevance of the established biophysical theory in your evaluation of the epidemiological data and the whole animal bioassays to your risk assessment.	prior to seeing any EMF-specific research results we thought that the energy associated with residential and occupational EMFs was so small as to be unlikely to have any significant effects. This pulled our prior down somewhat. And the biophysical arguments had a small impact on one of the reviewers priors..
Robert Goble	As noted above we believe that it would be worth making clear the difference between "theoretical modeling" and "laws of nature". The biophysical argument is often presented as reflecting the constraints of the second law of thermodynamics and any evaluator should be strongly influenced by violations of that. The critical argument is that what is really at issue is how well the simplified models reflect the underlying biological situation; and in that respect we share the skepticism of the evaluators.	We agree. We shall make this distinction clearer in the final document.
Ben Greenebaum	Models based on physics are generally simplified. They can be quite useful if the simplifications do not throw out crucial aspects of the situation and can be quite misleading if these aspects are overlooked. The only way one can decide which is the case is to look at the data that one trusts. Data that is good—good experiments, reproduced, etc., casts doubt on the model. Data that is marginal may become more suspect if the model contradicts it. Basically, I support the general idea used in the study.	We acknowledge your support for our point of view
Mark Israel	I do not agree. The large body of experimental data seeking evidence of a relationship between EMF and biological effects related to cancer causation is widely interpreted as failing to identify such a relationship. In addition, significant questions have been raised as to whether power frequency EMF at environmental levels are capable of imparting sufficient energy to have adverse effects on living organisms. Thus, the appropriate hypothesis based on experimental evidence is that power frequency EMF do not cause biological effects related to cancer, and the prior degree of confidence should be "vanishing" small. While evidence to the contrary should be considered, any such evidence must be interpreted in light of existing physical theory and the scientific evidence on which such theories are based. The decision not to be influenced by arguments based on physics or	We disagree with the statement that "The large body of experimental data seeking evidence of a relationship between EMF and biological effects related to cancer causation is widely interpreted as failing to identify such a relationship". All formal recent evaluation of these data (NIEHS, NRPB, IARC) have concluded that the risk is "possible". We do not question the physical theory, but the adequacy of the simplified biophysical models based on this theory. It is true that the use of "simplified models" is well established, but we believe that their purpose is to guide the experimental scientists, not to override their empirical findings. Whether these findings are strong or weak, consistent or inconsistent is a matter for the evaluation. It seems to us that, by requesting "extremely strong and consistently

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
	"simplified biological models" ignores that both of these approaches are well-established and are fundamental scientific tools for seeking evidence of causal relationships. In testing hypotheses, good science must take into account existing scientific knowledge. When a hypothesis (such as the hypothesis that power frequency EMF cause biological changes related to cancer) appears to be in conflict with the principles of physics, extremely strong and consistently reproducible experimental evidence is required to validate the hypothesis. Given that the experimental evidence in this instance is neither strong nor consistent, it is particularly inappropriate to simply bypass the question of physical plausibility in evaluating the hypothesis. Similarly, the use of model systems is a key means of eliminating confounding variables in evaluating potential causal relationships. Virtually all aspects of disease causation are routinely studied in model systems. For these reasons, an evaluation of EMF and cancer causation should include an integrated analysis of experimentation using model systems and the questions raised by application of the principles of physics. The lack of such an analysis is a significant omission in the draft CDHS Report.	reproducible experimental evidence is required to validate the hypothesis" the commenter is not entirely ruling out the hypothesis, that is, his prior is not vanishingly small, although it may be much smaller than ours.
Leeka Kheifets	My prior degree of confidence is very low (but I would not say it is vanishingly small). I do not believe that average fields of 3–4 mG can produce health effects. Nor do I believe that we can distinguish between 1 and 4 mG with existing methodology. However, that does not mean that I am inclined to discount the observed association with childhood leukemia based on simplified biophysical calculations. Rather, I believe we have a clue (currently best captured by average fields above 3–4 mG) that needs to be understood. Ideally, biophysicists will provide us with testable hypotheses, but even lacking these, their arguments are important in checking sanity of proposed theories.	Since the commenter states that the current evidence "needs to be understood", we conclude that she does not dismiss the hypothesis pureley on the strength of the physical models alone and that therefore she does not disagree with our position.
Patrick Levallois	Yes. It is possible that part of physics is not enough developed to be able to explain possible effects of EMF in humans	We acknowledge your support for our point of view
David McCormick	Several physicists have made the theoretical argument that the energy contained in 50/60 Hz magnetic fields is insufficient to induce biological effects of any type. Although this argument is attractive in its simplicity and absolute nature, it is not supported by empirical evidence gained from studies of signal amplification in well-studied biological systems. Perhaps the best	We acknowledge your support for our point of view. We agree that biological effect and health risk are two distinct concepts. However, the plausibility of the latter depends on the possibility of the former.

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
	example of this process occurs in the vertebrate retina. Experimental studies have demonstrated that one photon of red light contains approximately 3 x 10^{-19} joules of radiant energy. Capture of a single photon of light by a vertebrate photoreceptor cell produces a receptor current of approximately 5 x 10^{-14} joules; thus, the energy contained in the visible light has been amplified by a factor of more than 10^5 by the biological system. Should a comparable amplification process be involved in EMF reception by the cell, the energy delivered to a biological system as a result of EMF exposure could greatly exceed the amount of energy contained in the incoming stimulus. It is important to note that our understanding of sensory receptor systems provides a possible mechanism through which the energy contained in power frequency EMF may be amplified to a level at which biological effects may occur. However, no such amplification mechanism for EMF has been identified. As such, any discussion of possible amplification mechanisms should not be interpreted to support the contention that EMF exposure has biological effects. Rather, this issue is raised to indicate that the possibility of biological activity cannot be excluded solely on the basis of the physical arguments that have been put forth. A second, even more important caveat to this comment is that the induction of biological effects is in no way equivalent to the induction of adverse biological effects. Restated, the fact that biological effects of any chemical or physical agent can occur (or do occur) cannot be construed as evidence of human health risk.	
Thomas McKone	In my comments on draft two I supported this perspective and expressed the view that the discussion is sound and provides adequate justification for the position. My earlier comments still apply to draft three.	We acknowledge your support for our point of view
Samuel Milham	I agree. Theories which arise can often be tested with available data. Epidemiologic evidence should stand until refuted by better studies.	We acknowledge your support for our point of view
Hal Morgenstern	I agree that the prior should not be "vanishing small," but not necessarily because "theories should be used to predict results that are falsifiable and should not be used to discount evidence." How does one falsify the null hypothesis if that is what the theory predicts? See also (b) above.	The collection of robust empirical evidence of an effect falsifies the null theory.

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
Herbert Needleman	Good data trumps theory every time. I agree that the null assertions by physicists have little authority.	We acknowledge your support for our point of view
David Ozonoff	I agree completely with the approach that claims of "impossibility" or "implausibility" based on underlying physical or biological models should not be determinative or even greatly influence the evaluation of the evidence. I think many of the specific reasons given in the text are completely valid and I endorse them. I do not agree with the more general reason given in your specific question that theories should be used to predict results that are falsifiable. This idea (that science is "demarcated" by its production of falsifiable statements) has long been abandoned by philosophers of science (it belongs to the last time there was anything even approaching consensus in the field which was more than 40 years ago) and the text of the Review makes clear why it doesn't work: "falsifiable" predictions made by the theory do not really falsify anything. The existence of EMF bioeffects would not falsify electromagnetic theory, only suggest that some assumption or background condition is incorrect. We need to get by the "falsification" canard in epidemiology.	We acknowledge your general support for our point of view, while noting your caveat. We do not expect that empirical data of EMF effects will falsify electromagnetic theory, but only the adequacy of theorethical models.
Charles Quesenberry	The position taken by the reviewers of not being greatly influenced by arguments for zero probability of EMF effects at residential and occupational levels seems reasonable, and adequately justified in this document.	We acknowledge your support for our point of view
Jack Sahl	As supported in my main comments, you do not appreciate the value of the biophysical data to evaluate risks from a physical agent. You need to add expertise to your review team in order to correct the error in the draft.	Our question does not refer to data, but purely to "a priori" theoretical modeling, which, according to some, dismiss the hypothesis altogether.
Rick Saunders	I think that, when making a judgement on the plausibility of an environmental agent acting as a possible carcinogen, you should take all the evidence into account and consider the physical plausibility, the strength of the biological evidence as well as the epidemiological evidence, particularly when relative risks are small and the possibility of selection bias, residual confounding and chance may provide an alternative explanation.	We did consider biological plausibility both in setting our prior (which is the topic of this question) and in evaluating the evidence. The question, which is not answered by the commenter, is whether the plausibility is so low that the hypothesis should be dismissed a priori.
David Savitz	I agree that theoretical arguments do not explain away data, but I would approach the overall assessment of causality by integrating the epidemiologic evidence (with its strengths and limitations) with the experimental and	We agree with this comment. We believe that we have attempted to integrate all streams of evidence. However, some of us regarded the epidemiological evidence strong enough to override the lack of a

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
	theoretical arguments from other disciplines into an integrated evaluation. If the epidemiologic evidence were strong enough, the "hit" from theoretical counterarguments would be of little importance, whereas when the epidemiologic evidence is modest at best, an integrated assessment of the evidence would in fact give a fair amount of *relative* weight to the theories from physics. I also view the other disciplines not just as help in interpreting the epidemiology but also rather as independent lines of evidence, which, along with epidemiology, help to make an overall judgment.	biophysical mechanism. In any event, the comment indicates that the hypothesis should not be dismissed a priori, ie, the prior should not be vanishingly small.
Joachim Schüz	I fully agree with you that these theories should not discount epidemiological evidence. When it comes to causality, however, the lack of a biological mechanism is an important issue.	We acknowledge your support for our point of view as far as setting the prior. We did consider the question of mechanism in deriving the posterior.
Asher Sheppard	I am not in full agreement. In abstract terms, experimental data are more valuable evidence than theoretical analyses. However, where the data are highly uncertain, the question of whether to trust shadowy data can be helpfully answered by use of theory. If, as in the noise-based analyses of weak EMF fields, theory shows a gap of several orders of magnitude between the EMF level believed to be effective in some research study and a theoretical minimum, I find the theoretical analysis is a powerful argument against weak data. The frequently made appeal to unknown (and perhaps unknowable) factors attributed to biological complexity overlooks the very power of the biophysical argument, which is its simplicity. Specifically, there has to be some first interaction of a field with charges or moving charges and the magnitude of this interaction can be estimated without a sophisticated biophysical model. *The power of the biophysics argument is its simplicity and this simplicity is not a weakness as portrayed in the Risk Evaluation*. I acknowledge that any simple model may make a false assumption or overlook a key element. One such assumption may be equating 0.3 or 0.4 μT with the effective field in epidemiologic studies, when the effective field may be some other exposure feature highly correlated with exposures in the tail of the distribution function. Once the margin between theory and experiment or epidemiology is reduced to a small factor, then questions of biological complexity become relevant and much more interesting. In my opinion, there has been a degree of fuzzy thinking about the complexity of biological systems as if complexity were an answer to fundamental questions of information theory and energetics. Once the biological system receives a	This comment, although thoughtful, refers to the interpretation of data in the light of a theorethical model which tends to dismiss them as artefactual. It does not address our question, which is: should these models dismiss the hypothesis "a priori", that is before evidence is evaluated?

Question 1. We have taken the position that we are not greatly influenced by arguments based on physics and simplified biological models that suggested that residential and occupational levels of EMFs can't possibly produce bio effects. We say that theories should be used to predict results that are falsifiable and should not be used to discount evidence. Thus, our prior degree of confidence is not vanishing small. Do you agree? Please comment.		
	signal, biological complexity may be very significant, but biological transducers in the eye, ear, skin, and elsewhere are devices for movement of ions in a cell. It makes sense to examine mechanisms involving charge movement when investigating for a first cause of field interactions and to use simple models to do so.	
Gilles Theriault	I surely agree with this view. I congratulate you for having the courage of stating it openly. However, one remarkable feature of the entire research on the health effect of EMF has been the wide multidisciplinary approach taken to study it. Rarely in history has a health question been studied by so many diverse disciplines and by so many scientists of varied background in such a short period of time. This has lead to a very rich body of knowledge and has forced each discipline to go beyond its own limits in addressing the question and in scrutinizing the results. The confrontations of scientists' viewpoints on the same topic could only have lead to a collective enrichment that should be acknowledged in the history of research on EMF and health.	We acknowledge your support for our point of view
Lorenzo Tomatis	I agree that arguments based on physics and simplified biological models should not have a prevailing influence on your evaluations. However, the statement concerning falsifiable results sounds as popperian extremism. I would delete it.	We acknowledge your support for our point of view. We note your exception regarding "falsifiable results"
Jim Tucker	I agree that theories should not be used to discount evidence. I also agree that the prior degrees of confidence should not be zero. However, I feel that the prior degrees of confidence should be quite small, say 1-2%. This is based on my belief that our understanding of biophysical principles is quite good and that there is not likely to be a gap in our knowledge of biophysics that would leave room for EMF to cause biological effects.	We agree that the understanding of biophysical principle is quite good, although not complete. However, we feel that the models used to desribe these principles are not, nor can be, complete. Successive, improved versions of these models have yielded different thresholds for biological effects, suggesting that credible results are still elusive.
Nancy Wertheimer	A wise decision I think. The epidemiologic evidence combined with a small but growing number of positive laboratory findings not easily explained by present biological knowledge is now sufficiently strong to ask biologists to look beyond the present knowledge.	We acknowledge your support for our point of view

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
Name of respondent	**Response to question**	**Staff reply**
Theodor Abelin	The notion of initial degree of confidence is interesting and useful, because it aims at taking into account the tendency to accept harmfulness, which I guess is a personality characteristic. It is difficult to say what the personality characteristic really reflects, but it must have to do with mental flexibility and openness. The effect of 'a priori' expectations is comparable to that of theory (Question 1), because the more strongly a person leans toward a preconceived notion, the more resistant he or she may be to accept evidence as convincing. But looking at the three reviewers' reasoning, personality characteristics seemed less influential than scientific argumentation. Considerations explicitely (reviewers 1 and 2) or implicitely (reviewer 3) related to the evolutionary development seem particularly convincing to me. An important thought is that of reviewer 1 relating to the small number of persons likely to be affected. It relates to that of reviewer 2 noticing that electric and magnetic phenomena are involved in normal physiology and thus might on the one hand be tolerated, but on the other hand might interfere with it and lead to human pathology. A problem with this analysis of prior degree of confidence is that it had to be judged at a moment, when the reviewers most probably had already had some notions about observations of health effects. As far as I am concerned, it seems too difficult for me to rate myself quantitatively, and I will give a qualitative answer. I was very doubtful about any ill effect of EMF before I started working in this field, and I also conveyed this view to friends who told me about their own experience of exposure and its alleged effects. My own research in this field started because I was asked by a government agency to do so, and not really out of my own interest, and at that point I was still very doubtful about any biological EMF effect. Then, only when our data from several sub-studies (on RF-EMF and sleep quality) showed the same pattern, did I develop more openness toward this question. The literature on EMF and melatonin, and later on leukemia further influenced my position.	We admit that it is difficult to formulate a prior pretending not to know the existence of evidence. We tried to include in our reasoning only arguments that might have been defended without the use of information resulting from research specifically targeted to bioeffects of EMF.
Ross Adey	I disagree. There is at least one excellent study not cited in the Report relating residential electrification to the emergence of the childhood leukemia peak:	We need to make a distinction between the information that could have influenced our prior and the information that should affect our posterior.

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
	Milhain S, Osslander EM. Historical evidence that residential electrification caused the emergence of the childhood leukemia peak. *Medical Hypotheses* 56(3):290-295, 2001. The authors conclude, *inter alia,* that "During 1928-1932, in states with above 75% of residences served by electricity, leukemia increased with age for single years 0-4, while states with electrification levels below 75% showed a decreasing trend with age (P = 0.009). During 1949-1951, all states showed a peak in leukemia mortality at ages 2-4. At ages 0- 1, leukemia mortality was not related to electrification levels. At ages 24, there was a 24% (95% confidence interval, 84 1 %) increase in leukemia mortality for a 10% increase in the homes served by electricity. The childhood leukemia peak of common acute lymphoblastic leukemia may be attributable to electrification." These data should not be ignored. The position stated in your question appears to be a perpetuation of the ancient shibboleth perpetrated by the physicist Jackson at UC Berkeley in his 1991 paper published with the imprimatur of the National Academy of Sciences.	We were aware of the this article prior to its publication but could not use it for our evaluation because of the deadline for inclusion articles that we had set ourselves (publication by June 2000). Now that it has been raised in the comments, we hasten to respond. The Milham et al article points out articles from the early 1960's , that wee should have cited, in particular one by Court Brown and Doll (Leukemia in Chldhood and Young Adult Life, Trends in Mortality in Relation to Aetiology BMJ 1961:26: 981-986) These authors said : "..it is seen that the 2-4 year-old peak in childhood mortality was first recorded in in 1921-30 and that subsequently there was little change in the pattern of mortality until 1940-4, when the mortality at ages 2-4 years again began to rise relative to the mortality at younger ages. …. Among the white population of the U.S.A. the pattern of childhood mortality has been similar." Court Brown and Doll wonder if better treatment delayed death until age 2 and 3 or if sulfonamides might have saved children from infectious deaths long enough to be diagnosed with leukemia. Indeed one needs to ask if this age specific increase in leukemia mortality was real or an artifact of improved diagnosis. According to Shimkin (Shimkin M Contrary to Nature, US Department of Health and Welfare Publication No. NIH 79-720) adult leukemia was recognized simultaneously in the mid 1800s by Rudolph Virchow in Germany and John Hughes Bennett in Scotland on the basis of the naked eye inspection of the blood and spleen during autopsy. The blood looked yellowish white like pus. By the turn of the century microscopic examination of the blood would show various types of white blood cells. Myelocytic and lymphocytic types of leukemia were recognized. It is possible that the spread of routine microscopic blood tests after 1920 lead to the correct identification of leukemia, at first in the cities and then in the countryside. But let us, with Court Brown and Doll, accept the appearance of the childhood peak as real (since better treatments and sulfonamides were not present in 1920) and not an artifact of improving diagnostic procedure. Then for childhood

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
		leukemia the possibility that EMFs could convey a very high relative risk for the top 95th percentile of residential exposure should have been more credible a priori, since it would take a very high relative risk for the most highly exposed, to produce an average three fold increased risk for the population of toddlers over the range of exposures in an electrified environment. Our position is that Court Brown and Doll's (and a few earlier author's) observation of an emerging childhood peak, changes the possible size of the relative risk that we consider credible for childhood leukemia and not the a priori confidence that some EMF effect is possible. Milham's article is an EMF oriented piece of research and is relevant to our posterior degree of confidence. He shows that the peak in toddlers appeared first in places where the census reported widespread use of electricity. Thus something(s), (perhaps electricity) in those modernized locations caused the risk of toddler leukemia to increase three fold. The impact of the mystery factor(s) can be calculated (Population Attributable Risk Percent) and it would account for the majority (66%) of the leukemia in that age group. Now if the analytic epidemiological studies were coherent with the descriptive epidemiology they would reinforce our confidence. But they are not. The case control and cohort studies suggest a possible doubling of the risk above a 3 to 4 mG exposure. Three mG is something like the 99th percentile of exposure. Doubling the childhood leukemia risk in the highest1% of the population will not triple the overall rate of the entire population. To triple the average risk, a greater than 3 mG exposure would need to convey a relative risk of nearly 200! Alternatively the lower reaches of typical residential exposures would have to convey moderate risks. But the pooled analyses of the analytic epidemiology studies do not suggest either the hypothesis of extraordinarily high risk above 3 mG or moderate risks at low exposures. When presented with this anomaly, Milham (personal communication, 2000) invokes exposure

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
		misclassification in the analytic studies. But it is hard to believe that exposure misclassification would erase a relative risk of 200. Could exposure misclassification obscure moderate relative risks in lower zones of exposure common to most subjects? We have tried to hypothesize a series of relative risks from low to high exposure that would produce a three fold overall increase and then tried downward biases that would erase them all except for the very highest exposure. We couldn't come up with a reasonable scenario to reconcile Milham's results with those of the analytic studies. Milham and Ossiander's study warrants more detailed analysis in the United States and other countries, for example the peak occurred in Japan in later years. For now however we suspect that electricity per se does not totally explain Milham's data and since it is not coherent with the pooled analysis it does not increase our degree of confidence.
Anders Ahlbom	I cannot answer.	
David Bates	Long term changes in disease rates are hard to interpret, but there is evidence that childhood leukemia might be increasing slowly. The point should be made that this disease is almost certainly multifactorial in origin and requires a combination of circumstances for its initiation. The question is whether EMF exposure increases the risk.	We agree. Our point is that, by itself, a slow increase in the prevalence of a disease is not sufficient to increase our a priori belief that EMF exposure is a risk factor. It could influence our a priori confidence in an effect large enough to influence overall national rates.
Bowman	The justifications for the prior distributions are reasonable. However, the other prior distributions that the authors discard do exist among EMF researchers. If the goal of Bayesian inference is to capture the consensus of the scientific community, the implications of these other priors should also be reported.	Although we have consulted widely with colleagues and expert consultants, as discussed elsewhere in this document, we regarded that the evaluation was our own responsibility and both the prior and the posteriors reflect our best professional judgement.
Shan Cretin	This is not my area of expertise, but I think I would have had relatively lower a priori estimates for EMFs. I am influenced in this by the lack of biophysical models and also by the sense that EMFs are likely to be correlated with a number of other factors that could be the culprit even if there are real and enduring associations between EMFs and certain cancers.	We acknowledge these comments.
Carl Blackman	The reasons laid out by each of the three core reviewers to establish their initial (prior) degree of confidence that residential or occupational EMFs could produce	We are aware of Dr Blackman's work, but could not use it to

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
	relative risks of various sizes are cogent and convincing. My degree of confidence would increase as I gained laboratory experience testing the environmental/occupation agent. In the case of EMF, my degree of confidence is higher than any of the core reviewers because I have performed experiments with these specific EMFs since 1970. My research has found bio effects at very low intensities, e.g., my published work has shown effects caused by exposures as low as 0.2 mG and 10 V/m.	influence our prior.
John Dawsey *et al.*	Your method of risk assessment is neither scientifically sound nor defensible. You have wrapped yourself in Bayesian methods, without actually performing Bayesian analysis. First, the CDHS draft report did not follow the procedures outlined in the CDHS EMF Risk Evaluation Guidelines. Second, the risk assessment methods used are not considered standard practice for evaluation of potential public health risks. Third, the methods used are not useful for performing scientific risk assessments. CDHS should use established risk assessment methods. CDHS should also increase the number of authors by including scientists with expertise in the disciplines that are relevant to the available scientific information. This also will help to make the assessments more representative of the wider scientific community and to improve the relative weighting of data from the various scientific disciplines (e.g., laboratory experiments, whole animal bioassay, epidemiology, and biophysics).	We take notice of these comments. However, they do not address our questions: Do [the arguments we used to set our prior] seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?
Robert Goble	While we think the prior choices and their rationale are reasonable, we mostly want to note the following: it is an important innovation to disclose these for each evaluator; the final determinations are only moderately sensitive to these priors and that is a worthwhile finding; and that the critical sensitivity is the weight of evidence accorded to the various streams of information.	We acknowledge your support for our point of view
Ben Greenebaum	I tend to come at the prior estimates in a way similar to reviewers 1 and 2, though I would caution reviewer 1 that while no specific repair mechanism for EM fields would have evolved over time, no EM-specific damage mechanism would have evolved, either, making existing repair mechanisms for existing types of damage at least partially relevant. The stress protein work has some relevance here; the issue is the degree to which EM adds to other stresses to the point where EM tips the balance of an organism becoming unable to ignore the sum of the effects of its environment.	We acknowledge your support for our point of view, and we note your additional comments.

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
Leeka Kheifets	I believe eliciting prior degrees of confidence is a tricky business and I do not think informal solicitations, such as your question, provide useful and comparable information.	
Patrick Levallois	Yes, they seem really reasonable	We acknowledge your support for our point of view
David McCormick	I have some difference of opinion with the prior degree of confidence expressed by all three Researchers. Although I may not agree with their conclusions, I do understand and accept the rationale underlying the positions of Researchers Two and Three. However, the rationale underlying the conclusions of Researcher One appears to be less objective. For example, without any relevant supporting data, Researcher One appears to use different criteria to accept or reject the possible existence of certain types of effects; this bias skews this Researcher's prior degree of confidence. For example, Researcher One accepts as highly probable some types of phenomena that have not been proven to exist (*e.g.*, perturbance of organism function by EMF interference with electrical signals), while he rejects other unproven phenomena that would suggest opposite outcomes (*e.g.*, existence of a repair mechanism). My prior confidence would not enable me to accept either type of phenomenon. Before entering the EMF research field approximately a decade ago, I held absolutely no preconceived notion as to the magnitude of the hazard (if any) that EMF exposure posed to humans. At that time, EMF epidemiology data were conflicting; high quality animal data in bioelectromagnetics were virtually non-existent; and no experimental or epidemiologic database existed for similar agents. However, even without a known mechanism of action, the ubiquitous nature of human EMF exposure provided a compelling reason to conduct hazard assessments of such exposure. Then and now, I look at EMF hazard assessment as a problem in environmental toxicology whose evaluation must be approached by a "sum of the evidence" approach in which several types of information are integrated and synthesized. Lacking data to support one or more cellular, biochemical, or molecular mechanisms of action, empirical data from experimental and epidemiology studies must be given primacy in efforts to establish and quantify the risks, if any, that may be associated with EMF exposure. Ten years ago, such data were either absent or conflicting; at that time, I considered it nearly impossible to predict possible hazards with any degree of confidence. The relevant database in studies of EMF health effects is much more robust today, and	We explained that the prior was determined using only common sense argument and basic scientific knowledge. Reviewer 1 justifies his belief that perturbance of organism function by EMF interference with electrical signals is probable using the common sense observation that it is easier to damage a complex system by intrducing an unintended component than it is to repair the damage thus caused.

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
	supports a more effective evaluation of the possible hazards of EMF exposure.	
Thomas McKone	As was my comment on draft two, I have found that in draft three, the reasons for these judgments seem reasonable. Draft three has addressed my comment that draft two did not have a clear statement explaining whether the experts did or did not confer among themselves.	We acknowledge your support for our point of view
Samuel Milham	One dramatic change in disease rates after the introduction of electricity was the emergence of the mortality peak at age 4 of childhood leukemia. This has been shown to consist solely of common acute lymphoblastic leukemia. In the US, the urban to rural spread of electricity between 1920 and 1955 is strongly correllated state by state with residential electrification. I'm convinced that this type of leukemia is nearly completely attributable to EMF's. (see Milham and Ossiander, Med. Hypotheses (2001) 56(3),290-295.	Please see our response to Dr Adey
Hal Morgenstern	My priors would be similar to those of Reviewer 2 (see pp. 9, 29-30), and I tend to agree with the reasons given in the report. See also (b) above.	We acknowledge your support for our point of view
Herbert Needleman	I found the explanation for the priors persuasive. One does not often encounter an evolutionary explanation for an expected toxic response. It was very useful.	We acknowledge your support for our point of view
David Ozonoff	I agree (for the reasons given in the text) that animal bioassays should not be determinative here. The effect of the introduction of electricity on various rates, besides being subject to all the qualifications given in the text, represent a type of ecological design known to be subject to severe bias from confounding and effect modification, factors which are especially pertinent over the long time spans considered here. Thus such arguments are essentially useless (you can refer to the work of Morgenstern, Greenland, Robbins, etc., etc., here).	
Charles Quesenberry	I felt that the reviewers' final choices of the prior distribution are similar to what I would have chosen. I cannot recommend any changes. The development material in Chapter 2 is well presented, and useful in interpreting their decisions. The three sets of arguments leading to their choices are reasonable.	We acknowledge your support for our point of view
Jack Sahl	I do not understand (even though I am very familiar with the relevant scientific literature), or agree with your judgments, especially those of Reviewer 1 and 3. Nor am I as excited by the miscarriage results, since no other literature supports this (and much of it discounts these results) and the 'story' does not make sense (that a one-time exposure above 8 mG would be this bioactive, or detectable given	We take notice of these comments. However, they do not address our questions: Do [the arguments we used to set our prior] seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
	all of the other exposures that are commonly experienced). When I started work in this area in 1984, I was influenced by the work of Adey and his colleagues, Blackman and his colleagues, and the PMR studies on occupational groups. My initial prior would be around .25, and was increased by the Savitz and London results. EMF-RAPID failed to find experimental laboratory effects and the NTP bioassays were negative (Canadian, Japanese, and U.S.). Linet, McBride, and the English failed to support a childhood leukemia effect at 2 mG, but Greenland and Ahlbom found support for an effect at 3 – 4 mG when pooling (The Greenland and Ahlbom results do not impress me nearly as much as the three CDHS reviewers – I would put myself much closer to the 'discussion' in the manuscripts). The biophysics arguments got a lot stronger over the years, led by Weaver. Many tried to find a new biophysical mechanism, but they all failed. None of the experimental laboratory results withstood scientific scrutiny. My studies of electrical utility workers reassure me that the PMR studies were in error. The melatonin hypothesis hasn't gotten legs. The pattern has not gone well – better studies don't support the early work. The advocates increasingly rely on vague theories of causation and are unwilling to state clear exposure-disease models. The best scientists are leaving the field, and those who remain can't get funding from NIH, even thought there are has been $500 million in seed research funds. I am confused by the childhood leukemia pooled results, so am willing to stay open to possible effects. My posterior is about .15 (I could understand higher posterior's, but nothing like those expressed by the CDHS team).	agent? For EMFs? Why?
Rick Saunders	I would have taken a more critical view of the a priori degree of confidence that can be expressed with regard to the possibility of biologically significant interactions with environmental EMFs. The view that the very weak electric fields induced by these exposures, compared to endogenous electrical activity, might have significant effects and that we have had no time to evolve specific defence mechanisms (page 28) is somewhat speculative. The analogy with UVR (lines 33 and 34) is interesting. What one can say here is that, in contrast to EMF biology, there has been steady scientific progress in the understanding of UVR interactions with tissue and of the molecular basis for skin cancer induction and immune suppression over the last 10-20 years, even if its exact role in melanoma aetiology is not yet understood.	Not being based on specific EMF evidence, the *prior* is by necessity speculative. The fact that these man-made temporally and spatially coherent fields are qualitatively different from endogenous electrical activity was an imprtant consideration.
David Savitz	The uncertainty that exists at present makes a rather wide range of priors	In setting our prior, we deliberately ignored all evidence resulting

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
	reasonable, and those of the three reviewers certainly fall into that range. They are probably giving more credence to the possibility than most of those who look at this issue would (though I don't have empirical evidence to support that contention). I would incorporate the physics, biophysics, and experimental evidence into my prior and say "not very likely" to all health effects, with the epidemiologic evidence modifying that prior to varying degrees across the various outcomes.	from EMF research, using only common sense arguments and general scientific knowledge. In any event, our priors, which range in the few percent, could certainly be described by the words "not very likely" used by Dr Savitz.
Joachim Schûm	No comment. Actually, the part about the a priori confidence was the part I liked least in your report. But maybe I didn't understand the rational of it.	
Asher Sheppard	The priors given on p 29-30 (sec. 2.4) are reasonable and well argued, but unpersuasive individually and collectively. I was not swayed by the thought that environmental novelty and non-natural origins heightened the probability of harmfulness. The bounds chosen for undetectability and easy detectability of EMFs as harmful agents were reasonable. The argument (Researcher Three) that the EMF mixture might convey greater risk than an agent with one characteristic was not persuasive because it made the unfounded assumption of a probable risk from multiple adverse interactions. I mistrusted that these are indeed a priori judgments. Each of the three DHS scientists, and particularly Drs. Neutra and DelPizzo, has been prominent in the EMF research community for many years and has been weighing the pros and cons of EMF research continuously. How could they presume to have truly prior views, or did I miss something? As noted before, I am not fond of throwing probabilities about when they are essentially guesses and that applies here too.	We tried to anchor our prior confidence on the prior confidence that any chemical agent at residential exposure levels might produce an epidemiologically obsevable adverse effect. Then we adjusted our prior for EMFs from that, in most cases somewhat downward. Sheppards and several other commentors are not alone when they express discomfort with the subjectivity of stating a priori confidence . However we noticed that hidden assumptions about the prior credibility of the EMF hypothesis was important in the EMF debate. Hence we did our best to reveal our judgements on this matter to help decision makers see, from the ensuing debate the reasons why people disagree in this arena.
Gilles Theriault	The reasons underlying the initial (prior) degree of confidence by the three core reviewers are very reasonable. The three reviewers did not start with the same degree of conviction. At the end of the assessment, the ranking of conviction between the three had remained the same, but each one had seen his conviction gone up by some margin. This testifies to the quality of the assessment done. As for my own reaction, the reading of the report has increased my degree of conviction that the association between leukemia and EMF is not the result of biases or confounding or chance. In that sense, it has increased my belief in a causal relationship but I am still puzzled by the inconsistencies between studies as much on the exposure side as, and even more, on the disease side. The type of	We acknowledge your support for our point of view

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
	leukemia associated with EMF vary quite a bit between authors and are not necessarily reflected in the rate of overall leukemia. I will come back on this issue latter. Concerning the adult leukemia studies, I notice that the report lumps together the occupational studies with the residential studies. I think that the comparison with childhood studies [that are essentially residential in nature] should be made with residential adult studies. When one does this, one realizes that the residential adult studies yield results that are very comparable with the childhood studies on leukemia. I have published my views and analysis on these previously. [Theriault G, Li CY. Risks of leukaemia among residents close to high voltage transmission electric lines. *Occup Environ Med* 1997;54:625-628.] On other health outcomes, my a priori degree of conviction has remained quite the same after the reading of the report. Contrary to leukemia, the other health issues have not been studied enough to come to a definitive conclusion. I am impressed by the results of the studies on Amyotrophic Lateral Sclerosis (ALS) and await other original contributions.	
Lorenzo Tomatis	-----	
Jim Tucker	As noted in my answer to the first question, my prior degrees of confidence would be lower than each of the three reviewers, i.e. in the range of 1-2%.	We take note of this comment.
Nancy Wertheimer	(In line 6, change you're to your) My own prior degree of confidence in risk attributable to EMF exposure was extremely low at the time of our first report in 1979. It has increased a great deal since then, based largely on the weight of the cumulative epidemiologic evidence (which includes more than the occupational and residential studies: For instance a survey of electric blanket use and childhood leukemia might show considerable strength and consistency in its evidence of risk. Such use showed an O.R. of 1.5 in the Savitz study, 2.25 in the Linet data, and 7.0 in the London study (based on only 8 discordant pairs). The lack of a dramatic change since the introduction of electricity is not a concern to my mind, for several reasons including: (A) Milham's article with evidence that rural electrification did, in fact, introduce a new cause of leukemia in rural children. This paper should be cited in your	We take note of this comment

Question 2. Each of the three core reviewers have laid out their initial (prior) degree of confidence that residential or occupational EMFs could produce relative risks of various sizes. These estimates are constrained by what we know about animal bioassays for cancer and by the lack of dramatic change in disease rates after the introduction of electricity and as the use of electricity increased. Reasons are given for these judgements. Do they seem reasonable? How much higher or lower would your *a priori* degree of confidence be for any environmental/occupational agent? For EMFs? Why?		
	report. I attach a first sheet of the paper giving the reference. (B) My own knowledge, from years of field work, about how much prolonged exposure to magnetic fields has been decreased over the years by technologic changes in power delivery, etc., such as: (1) Use of non-conductive plumbing pipes and connectors. (2) Increased use of 240 volt (vs 120 volt) appliances (The latter can produce ground currents in house plumbing,; the former generally will not do so), (3) Installation of a plus and a minus live wire in the service to a house instead of a single live wire (the latter means all appliance-caused ground currents will add their effects; the former means appliance-generated currents can cancel each other, (4) Use of three-phase wires to deliver higher current loads, thus tending to mitigate any increase in fields due to increased loads, (5) Increasing the voltages used, thus allowing a smaller current to deliver a larger amount of power).	

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
Name of respondent	**Response to question**	**Staff reply**
Theodor Abelin	I agree. Mechanistic explanations seem to me to represent a vague form of theory (Question 1), whereas epidemiology provides evidence. Your domino-metaphor (page A-22/A-23) is very helpful in understanding this. In the past, I have used a black box metaphor, where epidemiology examines the association between inputs and outputs of a black box and leaves the study of the contents of the black box to subsequent research, which may be mainly on cell or animal models. The domino metaphor is more differentiated and therefore, more helpful. By the way, it has often seemed to me that laboratory scientists have found it easy to generate mechanistic explanations for their findings, whatever their observations may have been. Only, the resulting theories become less and less parsimonious. A last comment to this question (which may reveal that I myself tend to follow this pattern as well) is that I have been wondering, why your reviewers were not more positive about the possibility of a causal chain of EMF leading to changes of cell permeability (Ca^{++} efflux), reduction of melatonin excretion and loss of control over cell reproduction (i.e. malignancy). There seems to me to be some logic to this sequence, but I would agree that in view of present evidence, this may still be classified as speculative.	We acknowledge your support for our point of view
Ross Adey	I agree. To the extent possible, the epidemiology should stand as an independent body of evidence. Moreover, as in the case of cigarette smoking and lung cancer, public health policy should not await arrival of a complete body of scientific knowledge before establishment of essential health safety guidelines. However, there comes a point where further epidemiological studies are fruitless *unless they take full account of new and growing mechanistic knowledge*. We appear to be approaching that guideline in planning future research. The importance and validity of mechanistic knowledge is exemplified in the following example of the progression of new knowledge on the interaction of 60 Hz magnetic fields with inhibition of breast cancer cell growth by melatonin. The initial observations by Liburdy of inhibition of the melatonin antiproliferative action by 1.2 μT 60 Hz fields in 1993, has been confirmed and extended by three laboratories (Blackman et al., 1998; Luben et al., 1998; and Morris et al., 1998). The most recent study by Ishido et al., (Carcinogenesis 22(7): 1043 -8, 2001) confirms the previous work and provides *evidence for uncoupling of signal transduction from*	We acknowledge your support for our point of view. We agree that the replication of the original Liburdy experiments showing bio effects at intensities well below what some regard as the the "impossiblity" levels are important as bioeffects and even more important since they represent mechanisms that could conceivably lead to pathophysiological events. We will expand the discussion of mechanistic studies in that chapter to accommodate this comment and the articles it references. This tends to increase our degree of confidence somewhat

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
	melatonin receptors to adenylyl cyclase. Studies of this type reach to the very apotheosis of a mechanistic understanding. They give the lie to the epidemiologists' comfortable notions that it is appropriate to dismiss mechanistic studies as "insensitive and nonspecific."	
Anders Ahlbom	I agree fully	We acknowledge your support for our point of view
David Bates	I agree with this position. Lack of mechanistic understanding is not a reason to question the epidemiological results if these show consistency.	We acknowledge your support for our point of view
Carl Blackman	The core reviewers did not reduce their degree of confidence in epidemiology on the grounds that lack of mechanistic understanding is not sensitive or specific. I agree with this determination because EMFs are a complex assortment (cocktail) of possible causative agents and their individual influence on various components of cell biology leading to possible tumor formation is unknown.	We acknowledge your support for our point of view
Bowman	We agree that the theoretical and in vitro mechanistic studies do not explain how EMFs in the environment could cause the reported health effects. In most cases, the in vitro studies were conducted in model systems that are not remotely related to cancer studies, and are done with EMF exposures far higher than found in homes and most workplaces. Even if the in vitro studies showing EMF bioeffects are accepted as valid, they would not necessarily support a chain of events that might lead to pathology. We agree that the epidemiologic data should have more weight in the overall interpretation of the literature than these in vitro data that are unrelated to cancer development. This position is consistent with the other EMF reviews like the IARC framework where “limited evidence” from epidemiology and “inadequate evidence” from animals is sufficient to declare an agent a “possible carcinogen”. With EMF, the debate is mainly over the strength of the epidemiologic evidence.	As noted elsewhere in this document, the IARC conclusion that animal evidence is inadequate represented a majority viewpoint within the Working Group, not a consensus opinion. We agree that in vitro results showing EMF bioeffect unrelated to cancer development are not relevant if the exposure was above the 1 Gauss or so where everybody agrees that EMF bioeffect are possible. However, if exposure was below this level even if above the norma lenvornmental levels, credible experiments increase our confidence by refuting the argument that no effects occur below the threshold determined by theoretical models.
Shan Cretin	Perhaps I have already addressed this in Question 1. If there was a convincing mechanism, then the epidemiological studies should be more convincing and more interpretable. For example, if we know exposure is related to the cumulative dose over time or the presence of EMF lowering resistance to other carcinogens, then the epi studies could be interpreted in light of the appropriateness of the exposure measures and covariates. Without a mechanism to fall back on, it is harder to know what missing measures might explain away apparent effects or how poorly chosen exposure measures might fail to find true effects. So the lack of plausible causal	We acknowledge these comments. We acknowledge the limitations of epidemiological evidence, but we take the view that eviedence that can with stand scrutiny should not be weakened by lack of supporting evidence from other streams of research.

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
	mechanisms undermines my faith that the epi studies are appropriately designed, making me more apt to treat them as noise.	
John Dawsey *et al.*	No, we do not agree with your analysis. You have missed the key point. The important aspect of the available scientific literature is not that there is no established biophysical mechanism for the health risks suggested in the epidemiological literature (even though scientists have looked for such a mechanism for many years). Rather, it is that there are very well accepted biophysical mechanisms for the interaction of ELF/EMF and human cells. This is supported by the vast experimental literature and the results of numerous, and relevant, whole animal bioassays. The epidemiological data are less plausible given this available knowledge. In addition, your confidence in the epidemiology is misplaced. While the epidemiological literature can be described as 'limited,' we do not believe that, as scientists, you can confidently assert that we can rule out bias, confounding, or chance as plausible explanations for the observed associations in the pooled analysis (for exposures above 4 mG).	We do not question well accepted biophysical mechanisms for the interaction of ELF/EMF and human cells. However, we have no reason to believe that these are the ONLY mechanisms possible.
Robert Goble	We would perhaps put a slightly higher weight on the failure, after considerable attention to come up with plausible models. But we regard our disagreement as still lying within the appropriate range for the evaluators to exercise their judgement.	We take note of this comment
Ben Greenebaum	I agree with this point.	We acknowledge your support for our point of view
Mark Israel	After hundreds of studies over many years, the in vitro laboratory data fail to demonstrate that EMF is involved in the initiation, promotion or progression of cancer of any kind. Most importantly, despite extensive experimentation, the in vitro research has not shown that power frequency EMF is capable of causing damage to the genetic material of the cell (the DNA and chromosomes) that is known to be necessary to produce cancer. While the draft CDHS Report correctly notes that "there is no consistent pattern supporting genotoxicity" and "there is overwhelming negative evidence against DNA damage and chromosomal effects," these statements are essentially cast aside by the subsequent conclusion that "overall the picture is mixed and does not affect our confidence level much." The draft CDHS Report's complete discounting of the importance of the in vitro research is unwarranted. Strikingly, the draft CDHS Report's chapter on in vitro research is only 2 pages long. The brief discussion of research in this chapter does not identify even a single study relevant to the molecular biology of cancer. By failing to address this	The absence of evidence that power frequency EMF is capable of causing damage to the genetic material of the cell (the DNA and chromosomes) and our acknowledgement that "there is no consistent pattern supporting genotoxicity" and "there is overwhelming negative evidence against DNA damage and chromosomal effects," does not rule outthe possibility that EMF may play a role in the development of a cancer through non-genotoxic mechanisms.. It is not unprecedent to accept an agent as a cancer risk before a mechanism is discovered. Note that the IARC guidelines do not require a mechanism for a Group 1 classification. We did not describe the many null mechanistic studies at length because in our guidelines we had agreed not to

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
	important body of research in any meaningful way, the draft CDHS Report leaves the impression that either 1) there is no significant body of in vitro research, which is at best misleading, or 2) the in vitro research is not important in the overall evaluation of cancer causation, which is scientifically unjustifiable. The draft CDHS Report's discounting of the in vitro research is contrary to the approach taken by the National Institute of Environmental Health Sciences (NIEHS) in directing the recently completed six-year, $46 million national EMF research program (EMF-RAPID). As noted in the 1999 NIEHS Director's Report to Congress, a "major focus" of the EMF RAPID Program "was research that targeted examination of in vitro effects that might clarify potential mechanistic actions of ELF-EMF in order to explain reported epidemiological associations with magnetic fields." NIEHS emphasizes that the EMF-RAPID program focused on this type of controlled laboratory research because this research is an important element in the evaluation of cancer causality. Thus, as the NIEHS Director's Report made clear: [t]he NIEHS health effects research program focused on mechanistic, cellular and laboratory studies in the areas of Neurophysiology, behavior, reproduction, development, cellular research, genetic research, cancer and melatonin. … Mechanistic, cellular and laboratory studies are part of the overall criteria used to determine causality in interpreting epidemiological studies. (emphasis added) In reviewing the large body of in vitro research, the NIEHS Director's Report concludes that "most of the mechanistic work done in cells fail to support a causal relationship between exposure to ELF-EMF at environmental levels and changes in biological function or disease status." NIEHS considers this lack of evidence in the in vitro research a factor which "severely complicates" the interpretation of the epidemiologic research. The National Academy of Sciences (NAS) has also reviewed the results of the EMF-RAPID program. A 1999 report from the NAS National Research Council noted: When the EMF-RAPID program began, emphasis was placed on two important phenomena – cancer promotion and gene-related effects in vitro. Experiments supported by EMF-RAPID provided some evidence to support, and considerable evidence to refute the view that power-frequency MFs can have biologic effects. Evidence of any robust and replicated effects on the	duplicate previous efforts. Instead we refer the reader to the NIEHS working group report. The fact that most of the studies (with the exception of those cited by Adey in his comments above) were null, did not pull our confidence down so much, because we view these studies of elements of a complex mixture as prone to false negatives. In fact the NIEHS considered residential levels of EMFS as possible carcinogens for childhood leukemia and adult lymphocytic leukemia despite the null mechanistic studies. The biophysical theorists demand that experiments using level of exposures close to those in residential settings before they will believe the epidemiology, is a very stringent requirement that most other agents could not meet. Yet we have tended to accept this stricter requirement in our evaluation of the evidence, and have not given weight to mechanisms that are affected by magnetic fields in the thousands and millions of milligauss.

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
	development of cancer is lacking. (emphasis added) Rather than ignoring the in vitro research, the NAS Report concludes that the failure to demonstrate that EMF cause biological effects related to cancer causation is evidence against cancer causation. Thus, NAS concludes that "in view of the negative outcomes of EMF-RAPID replication studies, it now appears even less likely that EMFs in the normal domestic or occupational environment produce important health effects, including cancer." In my experience as a cancer researcher and as a past member of the Scientific Advisory Board of the National Cancer Institute, the approach taken by the NIEHS and NAS to include analysis of in vitro research is standard practice. For the authors of the draft CDHS Report to treat this research as lightly as they have is a significant shortcoming and compromises the opportunity for CDHS to provide a complete and accurate assessment of the body of EMF research relevant to cancer.	
Leeka Kheifets	While lack of mechanistic understanding tremendously hinders progress in this field, for me it is not a sufficient argument.	We acknowledge your support for our point of view
Patrick Levallois	No, there has been a lot of mechanistic studies and their mostly negative findings should influence more the confidence on the causality of the epidemiology findings	We take note of this comment
David McCormick	In the general case, the lack of a plausible biological mechanism is clearly not sufficient to discount a consistent pattern of epidemiologic and/or experimental findings of biological effects. It is important to note, however, that such findings would very clearly be strengthened by the identification of an underlying mechanism. What is absent in the present case is a consistent pattern of epidemiologic or experimental data that support the existence of a significant hazard of EMF exposure. In the case of EMF (and other agents for which the epidemiology database is inconsistent or equivocal), the absence of a plausible biological mechanism does indeed weaken the strength of the argument. Furthermore, the absence of confirming *in vivo* data from animal model systems also undermines the strength of arguments that are based on epidemiology alone. Absent either a plausible biological mechanism or supporting experimental data, my confidence in the modest effects identified in most EMF epidemiology studies is decreased.	We acknowledge your support for our point of view. We too regard mechanistic evidence a s a "strengthening only" type of evidence. We note your comments regarding the perceived weakness of the epidemioogical evidence. We deal with these comment elsewhere.

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
Thomas McKone	As noted in my comments on draft 2, I find this stance reasonable.	We acknowledge your support for our point of view
Samuel Milham	I agree. My favorite mechanistic theory has yet to be tested. There is some evidence that EMF's cause a failure of some parts of the immune system (in the U of W germ free rat study (Kunz et al, and in my study of healthy aluminum workers (Davis and Milham . (1990) AJIM 18, 79-85, immune cell phenotypes were similarly altered). Immune system failure could explain much of the EMF related pathology.	We acknowledge your support for our point of view
Hal Morgenstern	I agree that lack of a mechanistic understanding of how EMFs could cause biological effects should not pull down very much our degree of confidence in a (causal) effect. My reason is that historically many associations that were originally regarded as biologically implausible were later found to reflect true effects; moreover, lack of a mechanistic understanding for EMF effects may simply represent our etiologic ignorance.	We acknowledge your support for our point of view
Herbert Needleman	Lack of mechanistic understanding should not inhibit judgments about the reality of an effect, witness John Snow. We do not know the mechanism for lead's impact on the brain, but we know it happens.	We acknowledge your support for our point of view
David Ozonoff	Again, I agree completely and strongly with the approach taken here. The influence of knowledge of mechanisms is (as said) asymmetrical. Indeed it is my view that if and when such a convincing mechanism is produced there will be essentially an end to any EMF "controversy." On the other hand, if such a mechanism is not produced, the controversy will continue, unabated. This is validation of the approach taken here.	We acknowledge your support for our point of view
Regula Rapp	I am not convinced that knowing mechanistic explanations for EMF-effects are absolutely necessary before making policy decisions, or even before judging about causality. There are several examples of excellent epidemiologic evidence for causality before evidence from experimental studies existed (e.g. the famous Cholera epidemic in London, where Dr. John Snow removed the pump handle in Broad street and stopped the epidemic, suspecting dirty water as cause without knowing even the existence of bacteria).	We acknowledge your support for our point of view
Charles Quesenberry	My review of Chapter 4.7 of the NIEHS Working Group Report leads me to agree with the position that convincing mechanistic explanations do not currently exist. However, your arguments as presented in Chapters 4 and 5, and elsewhere, supporting your position that this does not greatly influence your level of confidence in EMF health effects are generally well reasoned and convincing to this panel	We acknowledge your support for our point of view

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
	member.	
Jack Sahl	They tried very hard, yet were not able to find a mechanism. This should count for something.	Again, we note that in many other instances the existance of a risk was accepted before a mechanism was identified. We are sure that in each case it could be said that "they tried very hard" to find one.
Rick Saunders	With regard to other known carcinogens for which supporting animal evidence was initially lacking, the large relative risks for smoking (page 8, line 51) and the existence of a clear dose-response relationship gave strong support even in the absence of these data. The difficulty I have is with the interpretation of small increases in relative risk in the absence of support from biological evidence. With regard to childhood leukaemias and EMF, AGNIR (2001) came to the conclusion that "In the absence of clear evidence of a carcinogenic effect in adults, or of a plausible explanation from experiments on animals or isolated cells, the epidemiological evidence is currently not strong enough to justify a firm conclusion that such fields cause leukaemia in children." We did add, however, "Unless, however, further research indicates that the finding is due to chance or some currently unrecognised artefact, the possibility remains that intense and prolonged exposures to magnetic fields can increase the risk of leukaemia in children."	We see this comment as essentially supporting our point of view. Certainly, it is easier to accept the epidemiological evidence if the risk is large. Whether the epidemiological evidence is currently strong enough to justify a firm conclusion that such fields cause leukaemia in children is the main focu of this evaluation and there is no unanimity between our reviewers.
David Savitz	See points above.	
Joachim Schüz	Once again, I think that the lack of mechanistic or animal data on adverse effects of EMF does not pull the degree of confidence in the epidemiological data down. It may be possible that the true mechanisms are more complex than those that we have examined so far. I'm not very convinced that one of the current hypothesis is actually the true mechanism, neither the melatonin-hypothesis, the contact current-hypothesis, or the transients-hypothesis. Nevertheless, epidemiology is only one piece of the puzzle. Moreover, it is not the case that the results of these epidemiological studies are irrefutable. It is also not the case that there is a lack of experimental evidence because they were no experiments; there were numerous studies on effects and mechanisms, but as you correctly said, the results were largely null. So I still think this decreases the degree of confidence in causality.	We take note of this comment. We acknowledge that non-supporting evidence exists, but we argue that the lack of positive results may be just as much reflect the limitation of our knowledge and our ability to plan and conduct adequate research, as the absence of an effect.
Asher Sheppard	When considering epidemiologic data, I agree that the lack of a biological mechanism is not a strong reason to discount the data. The absence of mechanism	We acknowledge your support for our point of view

Question 3. We were not deeply convinced of mechanistic explanations of how EMFs could cause bioeffects nor were we convinced of a chain of events that led to pahtology. Yet we did not let this pull our degree of confidence in the epidemiology down much on the grounds that lack of mechanistic understanding is not sensitive or specific. Do you agree? Please comment.		
	is good reason to check and check again, to look for confounders, biases, and so forth, as has happened for the childhood leukemia data and to lesser extents for other endpoints. While not ignorable, it is not a factor that can prevail against a large, thoroughly examined, set of data obtained with reasonably good techniques and decent exposure assessment.	
Gilles Theriault	I fully concur with the view that the availability of a mechanism of action or a chain of events that leads to pathology would contribute strongly to recognizing a causal relationship between some health outcomes and EMF exposure, but their absence do not negate the evidence as seen by epidemiologists and can only serve as a stimulus to keep searching for such a pathological mechanism.	We acknowledge your support for our point of view
Lorenzo Tomatis	I agree that the lack of understanding of the possible mechanism of action of EMF should not influence negatively a priori your confidence in the epidemiological findings.	We acknowledge your support for our point of view
Jim Tucker	No, I do not agree that mechanistic effects of EMF should not reduce confidence in biological outcomes. To me it seems quite inappropriate to allow some scientific arguments to be persuasive but not others. The biophysical arguments must be given their full weight and not excluded from consideration. To do otherwise is to instill a bias in the analyses.	We did not exclude them from consideration. However, even after careful consideration we were not convinced that the research conducted so far proves not only that a mechanism as not been found, but that no such mechanism may exist. This may weel be an impossible task. This is why we regard this type of evidence as “strengthening only”.
Nancy Wertheimer	I agree. No comment.	We acknowledge your support for our point of view

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.

IMPORTANT NOTE: We realize that this question is ill phrased. None of us thought that negative animal studies should increase the level of confidence. In fact this question should be restated as follows:

Question 4a: We were not swayed by the generally negative results of the animal toxicology studies, because we believe that they are prone to false negatives. Do you agree?

Many commenters agreed with our position that animal toxicology studies, following a protocol designed to test chemical agent, can be prone to false negatives when applied to a complex physical agent. On the other hand, many were not convinced by our arguments. As we have often noted in the course of this evaluation, we accept the fact that complex evidence may lead objective reviewer to different conclusions. We therefore respect these commenters' opinion. However they did not offer compelling arguments to make us change our conclusions. In summary our interpretation of the animal toxicology studies is this:

a) The epidemiological evidence suggests that there is no linear relationship between exposure and health effect (otherwise strongly exposed groups would exhibit risk high enough to be clearly visible). This suggestion is consistent with some of the proposed mechanisms (e.g., depressed melatonin production or any effect which reduces defense mechanisms rather than actively increase the health risk). Therefore, we do not accept that exposure to strong fields should result in proportionally higher risks that can be detected in a relatively small number of animals.

b) Rodents may not be an appropriate animal model. There are precedents in support of this possibility (e.g., the effect of smoking on lung cancer was only demonstrated in baby beagles, a most unusual animal model).

c) we believe that the EMF environment is far more complex than the simple sine wave fields used to expose test animals. However, we acknowledged (table 6.2.2, C2) that, if the Losher experiment are accepted a face value, then a sine wave should be also accepted as a valid exposure modality. In our opinion, the jury is still out on the Losher results, which remain unreplicated, but also unrefuted. We are not impressed by the observation that the rate of tumors in the controls in the attempted replications was typical for that type of experiments. When the cancer incidence rate in unexposed animals is higher than 90%, it is impossible to detect with any degree of confidence a modest increase in risk in the exposed animals.

d) Not all the standard toxicology tests are negative. This opinion was shared by the IARC working group, that defined the animal carcinogenicity data as "inadequate", rather than negative (majority opinion). A substantial minority characterized them as "limited" evidence of carcinogenicity. In particular, there is doubt on the significance of the thyroid cancer results (Boorman GA, McCormick DL, Findlay JC, Hailey JR, Gauger JR, Johnson TR, Kovatch RM, Sills RC, Haseman JK. Chronic toxicity/oncogenicity evaluation of 60 Hz (power frequency) magnetic fields in F344/N rats. Toxicol Pathol 1999 May-Jun;27(3):267-78), which, compared to past similar assays, suggest a very robust effect in one sex, but not in the other.

e) If the hypothesis that EMF is NOT an initiator is based on the results of studies reporting the absence of genotoxicity, DNA damage, gene expression changes, that Kavet accuses us of dismissing. These results are consistent with the epidemiological evidence suggesting a time lag between exposure and diagnosis inconsistent with the long latency period of cancer. However, proving that EMFs are not a cancer initiator does not weaken the hypothesis that they are a risk factor. We do not know whether EMFs are a promoter, a growth modifier or whether it acts to depress the body's defense mechanisms.

Most commenters were puzzled by the suggestion that a largely null body of work could increase the confidence level of reviewer 1. Reviewer 1 was including the chicken egg series of studies in this category, even though he viewed them as speaking to mechanism and the ability to produce an effect at intensities below the biphysical "impossibility" level, rather than as an indicator of mamalian reporductive pathology. This body of evidence will be presented under mechanistic studies and will affect the degree of confidence of that stream of evidence. (see Appendix- fpr a tab;e summarizing these studies.

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.

Name of respondent	Response to question	Staff reply
Theodor Abelin	To the extent that I have seen the literature on animal pathology, I agree with you. The animal studies I found were usually anecdotal, and observations in control animals were in general not reported. Therefore there would not be good reason to give much weight to the available reports. If you add that reports of animal pathology rather increased the degree of confidence, I interpret this as being based on the fact that whenever anecdotal studies are published they usually provide some level of evidence in favor of an effect rather than against. Is this, what you meant? Wouldn't this more likely simply express publication bias?	see above discussion
Ross Adey	I agree.	
Anders Ahlbom	I cannot answer	
David Bates	I agree with this position.	
Carl Blackman	The animal pathology literature is largely null effect, but the core reviewers did not let it reduce their degree of confidence in the epidemiological literature. I agree the animal literature should not have a negative impact on the epidemiology. In fact, it is possible to consider the animal literature, composed almost exclusively of exposure to sinusoidal fields of constant intensity, as supportive of more complex exposure paradigms which are closer to the actual agent in real life exposures captured by the epidemiology studies, whether it was measured or not.	
Bowman	We agree that the animal studies literature is largely null, and there is little there to suggest that EMF is a carcinogen. We do not agree that the breast cancer promotion studies or the chick embryo studies increase the degree of confidence in the epidemiological literature. All the studies reported by the Loescher group have limited internal consistency and therefore are not clear in identifying an EMF effect. This is particularly true in light of the Battelle failure to find similar results after an extensive effort to repeat these studies. Likewise, the overall literature on chick embryo studies is difficult to interpret because of the inconsistencies in the data across various studies. We do agree, however, with the interpretation that these data and the animal studies literature in general do not negate the findings of the epidemiologic data.	

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.		
Shan Cretin	The lack of animal effects does pull down my confidence in most human effects for reasons similar to those stated in 1 and 3 above. However, this would be a relatively small effect, since some diseases don't have easy to find animal models. The chick embryo studies *only* have *increased* my confidence for the possibility of effects on fetuses and very young children.	
John Dawsey *et al.*	You are alone in viewing the Loescher work as either relevant to cancer promotion or of value in addressing the question of potential health risks. The published reports from Loescher do not support any 'effect', the studies were not replicated, and there are results from other, well-designed and conducted studies that do not show any health effects. Your reliance on the 'Henhouse' studies is inappropriate. In 1997, a group of experts including two of the DHS reviewers unanimously concluded that the chick assay studies are equivocal and not a good assay for human risk assessment. The exposures for these studies are also not relevant to those found in community or occupational environments. You discount the lack of results from the majority (and best designed and conducted studies) of the animal bioassays by creating vague theories of disease causation. Even though these vague theories of causation were not addressed in the available literature, and you have no data to support them, you assume that the 'theory' supports the epidemiological literature.	
Robert Goble	And we have a similar view about the experimental studies. Though we would have put some of the evidence of biological effects up in the discussion of the "theoretical modeling".	
Ben Greenebaum	I would have given a bit more weight to some of the studies of behavior and of gene expression. Not much more, given the inconsistency between labs and the gap between in vitro/animal and human studies, but some. Epi got more discussion and more weight, as noted above, than other types of studies.	

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.		
Mark Israel	The draft Report's chapter on animals studies concludes that "[o]verall, the animal pathology studies are predominately, but not entirely, negative." As noted in Question 4 above, the conclusion that the animal research essentially provides no support for cancer causation, however, is essentially ignored or, more surprisingly, is used to increase rather than decrease the degree of confidence in the suggestive data from the epidemiologic studies. This dismissive approach to the animal data is entirely unjustified. Animal studies play an important role in the assessment of carcinogenic potential. These studies are used routinely in cancer research to assess the causative potential of many different sorts of agents and are a principal scientific method for determining carcinogenicity. The results of these controlled laboratory experiments should be given significant attention in any evaluation of cancer causality. The 1999 NIEHS Director's Report on EMF noted the importance of the animal research on EMF and carcinogenicity. > Animal carcinogenicity studies are routinely used to identify environmental agents that may increase cancer risk in humans. Many areas of biological investigation are more efficiently studied in animal models than in human beings. … The laboratory data in animal models are inadequate to conclude that exposure to ELF-EMF alters the rate or pattern of cancer. … [I]t is noteworthy that these data provide no support for the reported epidemiological findings (discussed earlier) of increased risk for leukemia from ELF-EMF exposure. (emphasis added) Similarly, the 1999 NAS Report on the EMF-RAPID Program concluded that: > The EMF-RAPID biologic research contributed little evidence to support the hypothesis that a link exists between MF and cancer. The results of the in vivo studies do not support an MF effect on cancer initiation, promotion, or progression, and they should be recognized as important studies in the overall evaluation of potential carcinogenic effects of MFs. (emphasis added). There is an extensive body of animal research that fails to demonstrate a detectable effect or role for EMF in the initiation, promotion, or progression of tumors. For this extensive body of research to be marginalized in the draft CDHS Report is inconsistent with meaningful scientific analysis. Perhaps more importantly, the failure of the draft CDHS Report to give this important data any meaningful weight in the overall causality analysis presents a misleading view of the research for the general public and regulators who may rely on the CDHS Report for information about EMF.	

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.		
Leeka Kheifets	I think you are giving too much weight to the breast cancer promotion studies (especially the Soviet results). Even more problematic, in my opinion, are the chick embryo experiments. I find that they provide no information at all. In the NIEHS evaluation, negative animal evidence pulled my assessment of the epidemiologic evidence down a bit. At IARC, I was persuaded by the argument that the animal data is not as negative as previously thought (based on the NTP experiment), although I did not think that there was enough evidence to classify the animal data as limited. Overall, I think high-quality and relevant animal data can be used to pull down epidemiologic evidence somewhat, but I do not see how it can increase the degree of confidence.	
Patrick Levallois	No, the mostly negative findings of the animal studies are a strong argument against causality. But I agree that this cannot dismiss totally the epidemiological results	

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.		
David McCormick	In my view, and in the view of the recently convened IARC EMF panel, the animal pathology literature is indeed null. However, the DHS Researchers appear to have accorded the results of animal studies relatively little import: according to the draft report, the DHS Researchers "did not let this pattern of [of experimental animal] evidence pull down our degree of confidence in the epidemiological literature," and "for some of us, it actually increased the degree of confidence somewhat." I find this conclusion to be an unfortunate situation in which epidemiologists focus solely on the data set with which they are most familiar and comfortable, and thereby ignore other evidence that could provide critical support to a hazard assessment. In cases where no clear pattern of risk (or lack thereof) emerges from epidemiology studies, the results of well-conducted experimental studies can provide important insight into the possible risks of agent exposure. A number of EMF epidemiology studies have identified a positive relationship between EMF exposure and the risk of a specific disease. However, studies in other, often comparable, populations have often not identified any relationship between EMF exposure and disease risk or outcome. As a result, few scientists would conclude that the sum of the EMF epidemiology literature provides a compelling case that EMF is a causal factor in human disease. In my opinion, the lack of experimental data to substantiate the positive findings of some EMF epidemiology should be considered in evaluations of the robustness of those findings. Although the vast majority of studies conducted in animal model systems have identified no hazards associated with EMF exposure, two types of studies were identified by the DHS as supporting a potential health hazard. For the reasons described, below, the results of these studies should be accorded limited importance. Studies performed using chick embryos as a model system are not commonly used to identify or evaluate human health hazards. As a result, the ability of such studies to predict human health effects is unknown: essentially no data exist to support the reliability and predictive nature of the chick embryo as an experimental model for human toxicity, and risk assessors and regulators do not commonly use chick embryo studies in developing human hazard assessments. The significant limitations of the rat mammary carcinogenesis data generated in both the Löscher laboratory and in Soviet Georgia suggest that the reported positive findings from these laboratories also merit only limited consideration. The value of the Georgian studies is limited by inadequate description of EMF exposure methods and monitoring, and by other concerns related to experimental conduct. As discussed in a previous peer-reviewed publication (Boorman *et al.*, Magnetic Fields and Mammary Cancer in Rodents: A Critical Review and Evaluation of Published Literature. Radiat. Res. 153, 617-626, 2000), the results of Löscher and colleagues	

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.		
Thomas McKone	In my comments on draft 2, I expressed the view that stance is reasonable and I sill hold this opinion.	
Samuel Milham	I agree. However, I think that the animal exposure studies done with other parts of the EMF spectrum should also be considered. The U of W rat study showed a large cancer excess in the microwave-exposed rats (18/100 cases vs 4/100 controls).	
Hal Morgenstern	I am not very familiar with the animal literature regarding EMF effects; thus, I do not feel qualified to judge how that evidence would affect my degree of confidence in causation.	
Herbert Needleman	Animal studies can support a theory, but cannot invalidate it, particularly if the number of subjects is 100 or 200.	
David Ozonoff	This is connected with the argument about mechanisms and is subject to the same reasoning. I strongly agree with the approach here.	
Charles Quesenberry	Similarly, upon review of the NIEHS report, your summaries in presented in Chapter 6, and appendix, your position with respect to influence of the animal pathology findings on your level of confidence is reasonable.	
Jack Sahl	The animal studies are very relevant to the evaluation of risk and should caution your exuberance about the epidemiology. My comments in the main section support this.	
Rick Saunders	I don't really see how a largely null animal pathology literature can increase confidence in the epidemiological literature. Experiments should be designed to tease out the biologically effective exposures, testing falsifiable hypotheses and maximising the chance of tracking down the aetiology of an effect. Sometimes, good animal models are lacking, or inappropriate exposures carried out. But I think that it would be wrong to conclude that the epidemiological evidence is strengthened on the grounds that the animal data were negative. To assume the existence of an appropriate environmental 'mix' that is more biologically effective than experimental exposures begs the question.	
David Savitz	See points above.	

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.		
Joachim Schüz	Once again, I think that the lack of mechanistic or animal data on adverse effects of EMF does not pull the degree of confidence in the epidemiological data down. It may be possible that the true mechanisms are more complex than those that we have examined so far. I'm not very convinced that one of the current hypothesis is actually the true mechanism, neither the melatonin-hypothesis, the contact current-hypothesis, or the transients-hypothesis. Nevertheless, epidemiology is only one piece of the puzzle. Moreover, it is not the case that the results of these epidemiological studies are irrefutable. It is also not the case that there is a lack of experimental evidence because they were no experiments; there were numerous studies on effects and mechanisms, but as you correctly said, the results were largely null. So I still think this decreases the degree of confidence in causality.	
Asher Sheppard	For much the same reason as above, the absence of strong evidence for animal pathology, while not ignorable, does not strongly affect causal inferences drawn from the epidemiology data. The fact of some positive data that have not been resolved (particularly mammary cancer in rats), inability to conduct studies at very high doses, and the limited statistical power of animal studies, augment my belief that the mostly negative animal pathology data are insufficient reason to significantly discount the epidemiology findings.	
Gilles Theriault	There are other examples where animal studies have contributed little to supporting known cause of diseases in human. The classical example is the inability to reproduce lung cancer among animal exposed to asbestos. A comparable example has been the generation of leukemia by exposing animals to benzene.	
Lorenzo Tomatis	I agree that the available animal data are of little significance, and I also agree that this should not pull down the confidence in the epidemiological findings. I would add that the lack of relevance of most of the animal data is due to the fact that most of the tests on EMF were conceived and carried out as if EMF were a chemical or chemical mixture. Different approaches may be, actually must be, explored.	

Question 4. We viewed the animal pathology literature as largely null, with the exception of the breast cancer promotion studies of the Soviets and Loescher's group and the various experiments with chick embryos. Once again because of argruments given we did not let this pattern of evidence pull down our degree of confidence in the epidemiological literature much and for some of us it actually increased the degree of confidence somewhat. Do you agree? Please comment.		
Jim Tucker	If "null" means that the body of evidence is ignored, then I agree, but only if the experiments were done poorly, in which case there is a scientific basis for their exclusion. However, if "null" means that the body of evidence is considered scientifically valid but that the results show no biological effect attributable to EMF, then these data need to be included in the analyses, just as do studies that claim an EMF effect.	
Nancy Wertheimer	I agree. Given our lack of understanding of EMF "dose", and of the host- and environmental-cofactors needed to produce effects, even the most carefully done animal work may be invalid.	

Question 5. Not all epidemiologists would agree with our position that relative risks between 1 and 2 should be taken seriously unless there is specified evidence for confounding or bias to explain it away. Do you agree? Please comment.		
Name of respondent	**Response to question**	**Staff reply**
Theodor Abelin	I agree with your statement that not all epidemiologists would be impressed by relative risks between 1 and 2, and I also agree with what seems to be your own position that an increase of risk by say, 20 to 60 percent should be taken seriously, if the studies have been done properly. The best known example of a risk increase of 20 to 30% is for passive smoking and lung cancer, and it seems to me that studies on EMF and childhood leukemia show even more variability in design, which, as you state correctly, increases the credibility of this association being causal.	We acknowledge your support for our point of view
Ross Adey	I agree. In this regard, I suggest that the Report does not appropriately consider, at least from a theoretical viewpoint, the fact that in civilized societies an unexposed population no longer exists. As is the case for ionizing radiation, there may be no threshold with respect to cumulative dose; and as Hoteling at UC Berkeley pointed out many years ago, low level exposures that are potentially pathogenic then become immersed in a sea of other low-level competing factors.	We acknowledge your support for our point of view. We also note your comment. We will add a discussion of this point. We do not have sufficient information to determine whether a non-exposed population exists. The childhood leukemia pooled analyses suggest, but not prove, that there is little or no risk below 4 mG. On the other hand, the two recent SAB studies suggest that common short exposure to strong fields are a risk factor.
Anders Ahlbom	I fully agree	We acknowledge your support for our point of view
David Bates	I agree that relative risks between I and 2 should be taken seriously. The demonstrated effect of particulate air pollution on respiratory and cardiac mortality provide a recent example where the force of the evidence is now being acknowledged.	We acknowledge your support for our point of view
Carl Blackman	Should relative risks between 1 and 2 be taken seriously? I believe they should because one does not know if the proper exposure metric is being tested in the particular study. It could be that the actual risk is generally low, or that a small fraction of the 'cases' are the only ones actually exposed to the active EMF components. Additional, focused research both in the laboratory and in the field is needed to establish probable metrics to evaluate.	We acknowledge your support for our point of view. We also agree that the risk could be stronger than what appears, because of exposure misclassification.
Bowman	I do not have enough experience with epidemiology to set a lower bound for reliable odds ratios. However, the relative risks of 1.2 from the meta-analyses on adult leukemia and brain cancer are not a scientifically reliable basis for a finding of possible carcinogenesis, especially with this wide-ranging collection of study designs and exposure assessments (see more comments on these meta-analyses	We agree that the risk estimate in those metaanalyses is not to be taken as an accurate estimate, because of the reasons you mention. However, the poor exposure assessment in those studies is expected to reduce the size of the association. Our evaluation is based more on the consistent

Question 5. Not all epidemiologists would agree with our position that relative risks between 1 and 2 should be taken seriously unless there is specified evidence for confounding or bias to explain it away. Do you agree? Please comment.		
	below).	pattern of results, as measured by the sign test, which does not depend on the size of the association.
Shan Cretin	Relative risks between 1 and 2 are not terribly convincing to me, especially when those risks are closer to 1 than to 2. I view the absence of a plausible explanation for confounding/bias as similar to the absence of a plausible biophysical mechanism. Surely there could be sources of confounding/bias not yet identified! Your assessment of these equivocal risks says that you are taking a conservative stance, in the sense of being biased towards finding an effect. As public health policy makers, this may not be an unreasonable approach but it is a bias.	We prefer to think that we are taking an "open mind" stance. When the possibility of bias cannot be ruled out, it is common to dismiss the evidence suggesting a risk as "not conclusive". We believe this is an appropriate position for an academic resercher, but does bias the evaluation by dismissing the possibilities that a) biases may NOT exist and that a RR of 1.2 represents a true 20% increase in risk, and b) that if they exist, they may bias the risk estimates downward, so that a RR of 1.2 may infact reflet a larger risk increase..
John Dawsey *et al.*	We disagree. We know from experience that there is a poor predictive value of epidemiological results for low estimated Relative Risks (e.g., review the contents of *American Journal of Epidemiology* or *Epidemiology* over the last ten years for studies that report estimated RR at these levels and note how the results are described). Your view is especially flawed in the context that there are small numbers of high exposed subjects and there is a lack of biophysical, experimental and animal support. With regard to small numbers, the pooled analysis by Ahlbom *et. al.,* reports that only 0.8% of subjects had exposure above 0.4 μT. The large majority of these subjects come from the study by Linet *et. al.,* who have demonstrated that participation bias and confounding occur in this study. In addition, no specific exposure parameter has been identified. In such cases, it is inappropriate to over-interpret the epidemiology.	We note your comment. Our criteria to interpret the epidemiological evidence is addressed elsewhere. We do not believe that past experience is very helpful. Obviously, past experience shows that strong pidemiological associations are likely to reflects real risks. It stands to reason that the first risks factors to be discovered are the strong ones. However, we know that the risk factors so far identified explain only a fraction, often a small fraction, of all the cases of cancers and other diseases. Many of the unidentified risks are likely to be intrinsically weak, but not necessarilly unimportant.
Robert Goble	Evaluating risks is a different exercise from constructing an epidemiological interpretation, and modest risks (as they appear in the environment) are the usual concern for risk evaluation. If epidemiological information is to be fully used in this enterprise, it should not be restricted to cases where there are large effects.	We acknowledge your support for our point of view
Ben Greenebaum	Some of this is a technical discussion among epidemiologists, but I think that the overall consistency of a significant number of good studies should be considered worthy of notice.	We acknowledge your support for our point of view
Leeka Kheifets	I find your question too general, size of relative risk is only but one consideration. Risks between 1 and 2 should be interpreted with caution, as they can be easily explained by bias and confounding. Nevertheless, for childhood leukemia, I do not think the existing evidence can be ignored, until we have specific evidence for bias	We agree with your comment. The reasons for our interpretation of weak epidemiological evidence is discussed elsewhere.

Question 5. Not all epidemiologists would agree with our position that relative risks between 1 and 2 should be taken seriously unless there is specified evidence for confounding or bias to explain it away. Do you agree? Please comment.		
	or confounding, or combination of both.	
Patrick Levallois	Yes, a RR of 1.5 can have a very important population impact. Unless evidence of bias or confounding, it should be considered seriously.	We acknowledge your support for our point of view
David McCormick	In the abstract mathematical sense, there is little doubt that a relative risk that is calculated to be between 1 and 2 may be real. Assuming that such an increase in relative risk is indeed real, the more important issues are (a) whether such a small increase in risk can be demonstrated with any degree of statistical certainty, and (b) whether an increase in risk of that magnitude is important to public health. In answer to part (b), it should be clear that in situations where even a modest elevation in risk is associated with broad population exposure, this small increase in relative risk could be an important determinant of disease incidence. However, demonstrating that such a small increase in risk is indeed real (and is not a function of confounding, bias, or random behavior within the study population) presents major challenges to study design and analysis. In this context, it is important to consider the limitations of epidemiology: whereas epidemiology is very good at identifying rare events superimposed on a background that is near zero (*e.g.*, angiosarcoma of the liver in workers exposed to vinyl chloride monomer) and high incidence events (*e.g.*, lung cancer in cigarette smokers), it is much less powerful at identifying small increases in incidence that are superimposed on a non-zero background. I will reiterate that in cases where the epidemiologic data are equivocal or conflicting, consideration of the results of studies conducted in experimental animals can be essential to the development of a hazard assessment. The results of the large body of high quality experimental studies that were designed to evaluate the risks of EMF exposure appear to have been accorded only limited importance by the DHS Researchers.	We agree with your comment. The reasons for our interpretation of weak epidemiological evidence is discussed elsewhere
Thomas McKone	I have no comment on this issue	
Samuel Milham	I agree. Most of the epi studies are of the case-control type, with an absense of clean controls. This gives low risks. With the fact that the entire population is exposed to EMF's, a small risk increase will effect large numbers of people.	We acknowledge your support for our point of view and accept the possibility that misclassification of exposure may mask a stronger risk.
Hal Morgenstern	Yes, I agree that estimated (?) relative risks between 1 and 2 should be taken seriously because: i) they might reflect a large impact on the disease in a population (i.e., attributable fraction) if the exposure is common; ii) they might be biased toward the null; and iii) they might reflect low frequencies of effect modifiers (e.g., due to biological interactions between EMF and other causes of the outcome).	We agree with your comment. In Draft 3 we also note that even for rare diseases a RR of 1.2 correesponds to lifetime added risks that are above regulatory “de minimis” benchmark risks.

Question 5. Not all epidemiologists would agree with our position that relative risks between 1 and 2 should be taken seriously unless there is specified evidence for confounding or bias to explain it away. Do you agree? Please comment.		
Herbert Needleman	I thought your explanation of the importance of small RR was correct. The consistency of positive reports is quite strong, and the population exposed is quite large.	We acknowledge your support for our point of view
David Ozonoff	With a highly prevalent exposure such as EMF it makes no sense (scientific or public health) to ignore "weak" effects. Indeed it is not agreed as to what constitutes a "weak" effect, since many investigators (e.g., Monson) consider effects above 1.5 moderate. I agree with this characterization, and moreover, as many texts emphasize, many causal effects are not large. In evaluating such effects one must (as was done) consider the likelihood that bias or chance produced the result. But this is taken into account with your method.	We acknowledge your support for our point of view. The prevalence of exposure to EMF fields > 3 mG is small in percentage terms, but since the whole population uses electricity, the actual number of subjects exposed is rather large.
Regula Rapp	I agree that relative risks between 1 and 2 should be taken seriously. Today, almost everybody is everywhere exposed to low frequency EMF, so you can't expect a wide variation in exposure. If you compare the risk of smoking 20 cig/d with the risk of smoking 25 cig/d, you will not get a big risk estimator, nevertheless the overall risk of smoking is high.	We acknowledge your support for our point of view and accept the possibility that misclassification of exposure may mask a stronger risk.
Charles Quesenberry	I do not believe that the position taken with respect to evidence for confounding or bias in interpreting relative risks less than 2.0 is overly problematic, particularly in the context of this evaluation of the various aspects of the stream of evidence from multiple studies for each endpoint under consideration - number of studies, consistency, dose-response, heterogeneity in populations studies, occupational vs residence.	We acknowledge your support for our point of view
Jack Sahl	It all depends on the strength of the available evidence. In this case, given the effort and the reassuring results from biophysics, experiments with cells and tissues, whole animal bioassay, and the epidemiology (which is not consistent on all stories), you should be very cautions about the epidemiology results.	We discuss the reason for our interpretation of the epidemiological evidence elsewhere in this document.
Rick Saunders	As for Q3.	
David Savitz	This is a key point – the lengthy, detailed consideration given to "what explains the positive epidemiologic findings" seems in my view to overstate the strength of findings that need to be explained. That is, with weak associations, inconsistent in magnitude, there's not a great need to invoke deterministic explanations based on causality or bias. Modest evidence for each or the relative merits of causality versus bias presuppose something about the strength of association to be explained.	We agree that the epidemiological evidence for some of the endpoints considered show weak associations, inconsistent in magnitude. However the magnitude of the association for childhood leukemia, to the extent that studies can be compared, is rather consistent, given the limitation of the measure. More important, the direction of the association is very consistent, as discussed elsewhere.

Question 5. Not all epidemiologists would agree with our position that relative risks between 1 and 2 should be taken seriously unless there is specified evidence for confounding or bias to explain it away. Do you agree? Please comment.		
Joachim Schüz	I agree with you that also small relative risks between 1 and 2 should be taken very seriously, if they arise from high-quality epidemiological studies. Nevertheless, one should be very cautious regarding bias when the risk increases are only moderate, and higher relative risks may be more convincing. For residential EMFs, one has to be especially cautious, because in most studies there are not only weak associations, but also the prevalences of exposure are very low (decreases the power of the study), the response rates are at most good-to-fair (selection bias is probable), and one looks at diseases for which little is know about the etiology (no idea about confounding or co-carcinogenity).	We acknowledge your support for our point of view
Asher Sheppard	Context is missing here and would influence the answer. Do you mean a case-control study? a large cohort study? a meta-analysis? To this non-epidemiologist, the dividing line here seems to be between those whose experience and observation of history show that small risk ratios rarely mature into confirmed causal relationships, and those who look at the data in isolation of history and focus on statistical features and study quality, among many other factors. I am sympathetic to the prejudice that arises from historical experience, but believe that such prejudices should be put aside until research on a topic has matured, for example as indicated when there are enough studies to do meta-analyses with fairly large numbers across the exposure distribution. I strongly favor letting the statistics tell the story even if the apparent RR is below 2.	We agree with your final statement. Please, see our response to Dawsey et al, regarding historical experience.
Gilles Theriault	Why would a study with a risk of 1.0 or even 0.8 been considered perfectly acceptable as showing the absence of a risk and a study with a risk of 1.5 or below been immediately judged as suspected of being flawed by biases and confounding. The size of the risk as nothing to do with the quality of a study. If a true risk is small, a well-done study will yield a small risk. If the study is powerful enough, this small risk will become 95% certain. I like the way you have handled this question in the argumentation under section 8, leukemia. The reply was that with so many studies from so many diverse populations and conducted by so many investigators, the same biases and confounding cannot have taken placed all the time and act in the same direction.	We acknowledge your support for our point of view
Lorenzo Tomatis	I agree in absolute that a RR between 1 and 2 should be taken seriously; in the specific case of EMF the existence of confounding or bias, although not entirely dismissed, does not seem to play a role. The precautionary principle suggests that evidence for a risk associated with an exposure to EMF should be taken very seriously.	We acknowledge your support for our point of view

Question 5. Not all epidemiologists would agree with our position that relative risks between 1 and 2 should be taken seriously unless there is specified evidence for confounding or bias to explain it away. Do you agree? Please comment.		
Jim Tucker	I would tend to agree that small relative risks should be taken somewhat seriously, providing that the confidence intervals are small enough to exclude 1. However, small relative risks should not be taken as seriously as larger risks. The extent to which relative risks are taken seriously should scale with the size of those risks. Similarly, the extent to which Society expends its public resources to solve a problem should scale with the magnitude of that problem. More resources should be used to address a relative risk of 2 than a relative risk of 1.2.	We acknowledge your support for our point of view and we agree with your comments. Our policy contractor's analyses consider the size of the association, the baseline incidence and the frequency of exposure as well as cost considerations in suggesting reasonable and proportionate response to the possible problem.
Nancy Wertheimer	I agree. Most epidemiologists aren't usually asked to think about epidemiologic evidence obtained where a "dose" can't be defined with any confidence -- and where the problem is made worse because we don't know if dosage is cumulative, or at what time prior to diagnosis the exposure should be evaluated. This is not a subject where traditional expectations based on cearcut independent variables and initiation of cancer can be applied.	We acknowledge your support for our point of view

Question 6. We said that a lack of specificity in the association of EMFs with subtypes of cancer and evidence for effects on various types of disease did not pull down our degree of confidence and might even increase our degree of confidence that epidemiological associations between disease X and EMF are causal in nature. Do you agree? Please comment.		
Name of respondent	**Response to question**	**Staff reply**
Theodor Abelin	Yes, I agree again. If there is a mechanism involving a reduction of anticarcinogenic activity (such as could be the case with a reduction of antioxidant activity), several if not all cancer subtypes could be affected. This is not comparable to local exposure to specific carcinogens such as those leading to lung or stomach cancer, or to tumor promotion by female hormones leading to gynecological cancers.	We acknowledge your support for our point of view
Ross Adey	I agree. But the question of EMF relationships with a variety of subtypes of cancer again raises essential questions about *mechanistic interventions.* At issue is whether EMFs may be a significant co-factor in aspects of carcinogenesis. Occupational exposures to high magnetic field levels along with metallic oxide fumes has been cited in occurrence of non-Hodgkin's lymphoma and immunosuppression in aluminum smelter workers (Davis and Milham, 1992). Weedicides and pesticides may be similarly involved, and in our domestic environments there are a myriad potential carcinogens.	We acknowledge your support for our point of view
Anders Ahlbom	I partly agree	
David Bates	I agree with this position; the risk in relation to diseases other than childhood leukemia is difficult to define, and the emphasis in this document is, I think, correct.	We acknowledge your support for our point of view
Carl Blackman	Should the lack of cancer subtype associated with EMF exposure and evidence for effects on various types of disease reduce or increase the degree of confidence that epidemiological associations between disease X and EMF are causal in nature? I believe we don't know enough about how a series of EMF-induced biological perturbations, and the possible biological states of sensitivity, can affect tumor formation. Without such understanding, we can only monitor changes that we think detrimental. Evidence that EMF exposure is associated (see comments on Scarfi et al. results below) with specific diseases indicates multiple consequences can occur from EMF-induced changes. The reviewers are correct to have this evidence increase their degree of confidence.	We acknowledge your support for our point of view
Bowman	The fact that the in vitro studies showing EMF effects involve primarily cell signaling and other epigenetic processes supports the interpretation that EMF might have effects that would be general on many disease processes and not specifically affect a target organ or disease. If these in vitro findings are confirmed and more robust evidence is developed to support them, the acceptance of associations with many	We agree with your comment and acknowledge your support for our point of view

Question 6. We said that a lack of specificity in the association of EMFs with subtypes of cancer and evidence for effects on various types of disease did not pull down our degree of confidence and might even increase our degree of confidence that epidemiological associations between disease X and EMF are causal in nature. Do you agree? Please comment.		
	different diseases is a reasonable position.	
Shan Cretin	The lack of specificity with regard to subtypes of cancers does weaken my view of the evidence. I'd certainly be more confident of an effect if faced with consistent, specific disease-exposure associations. I guess the main support for increasing confidence is if the posited mechanism is that EMFs somehow increase the "infectivity" of other agents. However, I was surprised by your argument here and do not agree.	We note that you do not view that lack of specificity as weakening your degree of confidence. In this important point, your opinion and ours coincide. One of the reasons for the lack of specificity increasing the level of confidence somewhat is that previous experience has shown that an agent that reaches many organs is likely to affect more than one health endpoint (eg, smoke is a risk factor for mouth, throath and lung cancer, UV is implicated in three different types of skin cancers, ionizing radiation increases the risk of many cancer types and subtypes.
John Dawsey *et al.*	We disagree. It is implausible that EMF is a 'general health hazard.' First the scientific data do not support this (e.g., neither laboratory experiments nor whole animal bioassays find robust suggestions for adverse effects on intact cells or tissues). Second, if EMF were a 'general health hazard,' this would imply that the disease model would be more conspicuous, which would suggest the whole animal bioassay and the laboratory experiments would find more robust results. This should either be neutral to your weight of evidence or diminish your confidence. Over the last thirty years, a vast number of exposures and different disease types have been evaluated. None of the earlier suggestions for an effect, including the '2 mG MF level' suggested by Wirtheimer and Leeper (1979), have held up to better studies. In contrast, Reviewer 1 uses this line of thinking to increase his belief that EMF is linked to health impacts. There is no evidence for a common biological model between the six diseases that Reviewer 1 concluded were likely to be caused by EMF exposure (i.e., childhood leukemia, adult leukemia, adult brain cancer, female breast cancer, spontaneous abortion and ALS). While Reviewer 1 concludes that he is virtually certain that EMF exposure is not a 'Universal Carcinogen', he does maintain that three fundamentally different cancer sites are linked to EMF exposure. A fair reading of the available scientific data does not support this.	The commenter misrepresent our question. There is difference between believing that EMF is a risk for more than one health condition and believing that it is a "general health hazard". We specifically state that the evidence ***does not*** support the propsition the EMF is a "general carcinogen", much less a general health hazard. In any event, the commenters position (that the lack of specificity "should either be neutral to your weight of evidence or diminish your confidence" is at least partly in agreement with ours. The only reason why it slightly increased Reviewer 1's confidence was that he felt magnetic fields were more likely a priori to affect many tissues and that c associations with several diseases weaken the argument that environmental fields have too little energy to affect biological processes.
Robert Goble	We agree with the arguments as presented, but want to note that they imply some general understandings about mechanisms; this suggests to us that mechanistic information, or the lack thereof, carries some weight.	We did not think that accepting lack of specificity implies some general understandings about mechanisms. It is possible that different organs react in different way to a common stimulous just as it is possible that the same underlying mechanism (eg,

Question 6. We said that a lack of specificity in the association of EMFs with subtypes of cancer and evidence for effects on various types of disease did not pull down our degree of confidence and might even increase our degree of confidence that epidemiological associations between disease X and EMF are causal in nature. Do you agree? Please comment.		
		a depression in the immune system) may be responsible for more than one adverse health outcome.
Ben Greenebaum	This was an interesting point and worthy of further discussion. I do not discard it out of hand. Are there analogous situations in chemical toxicology or malnutrition?	Ionizing radiation is a risk factor for many types of cancer. UV radiation is a risk for several skin health effects (erithema, premature aging, squamous cell carcinoma, basal cell carcinoma, malignant melanoma). The HIV is a risk factor for a variety of illnesses, including cancers.
Leeka Kheifets	Lack of specificity does pull down my confidence a bit, and certainly does not increase it.	We note this comment.
Patrick Levallois	Yes and No. Yes in general. But not really when you consider some specific illness as adult leukemia. As those sub-types of leukemia are in fact several kind of disease, environmental agent may act more specifically on a subtype of leukemia and not on all kind. I don't see how the observation of non specificity may increase the degree of confidence	As we have noted, we do not believe that EMF is a risk factor for all types of cancers or all subtypes of one cancer. The reasons why Reviewer 1 believes that lack of specificity may be considered an argument FOR causality are given in section 7.2 of the Draft Evaluation: 1. An agent that reaches EVERY organ of the body and that might affect signal transduction, and the endocrine and immune systems, would be highly unlikely to result in only one visible pathology. 2. It is much less plausible to believe that, among the factors correlated to EMF, there could be an unidentified risk factor for each of the many endpoints associated with EMF exposure. For example, it would be more plausible to believe that traffic fumes, which contain the leukemogen benzene, could be the cause of the association with leukemia, if leukemia were the only cancer associated with EMF exposure, but is not a credible confounder for breast cancer or ALS (Lou Gehrig's disease), which are not correlated to benzene. A similar argument can be made for bias. For example, it is less plausible to believe that selection bias in favor of higher versus lower social class could explain all the associations, because social class is a risk factor for

Question 6. We said that a lack of specificity in the association of EMFs with subtypes of cancer and evidence for effects on various types of disease did not pull down our degree of confidence and might even increase our degree of confidence that epidemiological associations between disease X and EMF are causal in nature. Do you agree? Please comment.		
		some diseases and a protective factor for others. All three reviewers believe that i the evidence for causality becomes convincing for one association, then the claim that environmental EMFs provide a dose too trivial to have epidemiologically detectable effects (which is one of the major arguments for a low prior) becomes moot, and the credibility of other association increases.
David McCormick	I would propose that lack of specificity with respect to disease site should most definitely decrease, rather than increase, the confidence that the DHS Researchers place in any observed epidemiologic associations between EMF exposure and cancer risk. It is unclear to me why this lack of specificity should increase anyone's level of confidence in the findings. Cancer is a family of diseases, rather than a single disease entity. Although neoplasms arising in different sites often demonstrate important similarities, different molecular alterations appear to underlie neoplastic transformation in different tissues. Furthermore, the kinetics, types of growth, responses to pharmacologic intervention, and other biological parameters often demonstrate a wide range of differences between sites. It is clear from both human and animal data that environmental agents demonstrate organ specificity in cancer induction; what is not clear from the cover letter to the Report is why our growing understanding of cancer biology and differences among tumor types should be not be integrated into EMF hazard assessments.	If we had a mechanism for EMF effect on one disease we could ask if it explained other disease associations. But we do NOT have a mechansims so we are left with the question whether specificity should have any weight a priori. See or response to Levallois.
Thomas McKone	In draft two, I found your stance reasonable and gave examples to support my finding. With regard to draft 3, I believe the stance is still reasonable.	
Samuel Milham	I agree. In my studies, there is evidence that some of the adult leukemia subtypes have an increased risk.	
Hal Morgenstern	I agree that lack of specificity in the association between EMFs and one disease should not reduce our confidence in causation for a given disease. Whether such lack of specificity *increases* our confidence in causation, however, is questionable. First, this increase in confidence depends on the plausibility of multiple or inter-related biological mechanisms. Second, similar associations with different diseases might be due to similar biases.	We note this comment. As explained in the Draft Evaluation (section 7.2) we do believe both that multiple mechanisms are possible (although no clear mechanism has been identified, several independent biologicsal effects have been tentatively identified) and that a common mechanism (eg, melatonin suppression) may be a risk factor for more than one endpoint). As for the possibility of a common bias, no credible candidate

Question 6. We said that a lack of specificity in the association of EMFs with subtypes of cancer and evidence for effects on various types of disease did not pull down our degree of confidence and might even increase our degree of confidence that epidemiological associations between disease X and EMF are causal in nature. Do you agree? Please comment.		
		exists.
Herbert Needeleman	I agree.	We acknowledge your support for our point of view
David Ozonoff	"Specificity" as an attribute of causal associations has long been disregarded by epidemiologists (see, for example Hill's own paper and many texts, e.g., the new edition of Rothman and Greenland which considers it essentially "useless.")	We acknowledge your support for our point of view
Charles Quesenberry	Your position with respect to lack of specificity in association is reasonable, and generally well argued in chapter seven. The argument for evidence regarding the association with one endpoint influencing the confidence in association with another endpoint is reasonable, primarily in the weakening of the basic argument of environmental exposures to EMF are at too low of a dose to result in health effects - as presented in chapter seven.	We acknowledge your support for our point of view
Rick Saunders	I would have thought that generally, specific agents are associated with specific types of disease for mechanistic and biological reasons. It is, however, easy to understand why different risk factors and diseases (but I think of largely viral origin) might be associated with something that depresses immune system responsiveness, such as AIDs (page 69), or immune suppression in transplant patients to give another example. I would suggest that experimental evidence for EMF effects on immune responsiveness is pretty weak. I think that it is wrong to conclude that the lack of association of EMFs with different subtypes of cancer increases confidence in epidemiological associations between disease X and EMFs.	We note your comment. Immune deficiency depression is only one example of an indirect effect that may result in an risk for more than one disease. Hormonal imbalance could be another.
David Savitz	Specificity or lack of specificity is not a major concern, except insofar as findings for one outcome help to form insights that are applicable to other outcomes.	We acknowledge your support for our point of view
Joachim Schüz	I agree that a lack of specificity should not pull down the degree of confidence in this field. The reason is that in most studies the exposure assessment relies on measures that may be biased by non-differential misclassification rather than differential misclassification. Lack of specificity may be an important issue in studies relying on recall by interviewees. However, it seems to me that there is some specificity in the association of EMFs with different types of cancer. To my mind, the childhood leukemia studies are much more convincing than studies for any other type of cancer. For adults, there also seem to be the strongest effects for hematological malignancies. I'm not	We agree with your comment. In the evaluation, we have specifically state that we see no evidence that EMF is a risk factor for all cancers.

Question 6. We said that a lack of specificity in the association of EMFs with subtypes of cancer and evidence for effects on various types of disease did not pull down our degree of confidence and might even increase our degree of confidence that epidemiological associations between disease X and EMF are causal in nature. Do you agree? Please comment.		
	convinced by the studies on brain cancer or breast cancer.	
Asher Sheppard	If EMFs are causally related to leukemia, brain cancer, and particularly subtypes of these cancers, we are so far from understanding the mechanism that I suspend any judgments that might be made from experience with chemicals and ionizing radiation. That is, the lack of specificity may be a feature of a hypothetical EMF mechanism about which we know so little that previous carcinogenic mechanisms are of dubious instructional value. However, from the pathologist's perspective on development of certain tumor types, this could be a facile and unintelligent remark. Therefore, dependent on details of the pathological nature of various tumor subtypes, the lack of specificity in cancer types generally is a weak argument for mistrusting epidemiologic findings. When applying this statement to EMF epidemiology, there is too little known about occurrence of cancer subtypes to make much use of pathology-derived views on etiology and therefore the apparent lack of tumor specificity is a still weaker argument against epidemiologic findings.	We agree with your comment
Gilles Theriault	I am more worry about the lack of specificity. This applies more acutely to leukemia. It is puzzling to realize that some studies observed increase in th risk of one leukemia sub-type and other studies in another sub-type, with or without an increase in the risk for overall leukemia. This lack of specificity pulls down my degree of confidence. I would need some reasonable explanation that I have not found so far. I think that the report does not discuss well the lack of specificity and the reviewers are wrong in using the lack of specificity as "even increase our degree of confidence that epidemiological associations between disease X and EMF are causal in nature".	By lack of specificity we were thinking of the range of effects ranging from ALS to Miscarriage to Cancer. The lack of consistency as to subtypes of leukemia between studies does bear discussion and we will enlarge the discussion of this.
Lorenzo Tomatis	That various types of cancer and various types of disease appear to be associated with exposure to EMF is plausible, but I would not take it, at least at present, as a reinforcement of the degree of confidence in the available epidemiological findings.	Not all reviewers did increase their degree of confidence because of lack of specificity and those who did, did so only weakly.
Jim Tucker	The lack of specificity between EMFs and subtypes of cancer could be argued as either reducing or increasing the degree of confidence of associations. In this case I'm not sure which is more appropriate, so I would tend to agree with the approach that was taken.	We acknowledge your support for our point of view
Nancy Wertheimer	I agree. In fact I expect lack of specificity, based on the likelihood that EMF is an enhancer of the carcinogenic process rather than an initiator.	We acknowledge your support for our point of view

Question 7. Have we done an adequate job in presenting the arguments for and against causality or are we assigning weak arguments to the "con" or the "pro" position?		
Name of respondent	**Response to question**	**Staff reply**
Theodor Abelin	Having read Chapter 8 on Epidemiology of the Leukemias in more detail than the other chapters, I respond with reference to this particular chapter. I think that you have done an outstanding job of reviewing each single criterion of causality (Chapter 8.2), and I have nothing to add. Your arguments against causality and for causality are well thought through, and your 'Comment and Summary' part is clear and equally convincing. Given my general and enthusiastic consent, I will not continue by making more detailed comments on particular chapters and lines of the text. I hope that this will be acceptable to you.	We acknowledge your support for our point of view
Ross Adey	I have given much thought to this question, because the ultimate evaluation of the historic value of the Report rides on the answer. I conclude that, in a clear attempt to avoid even a hint of reviewer bias, the preparers have failed to act as an independent committee of experts. I submit that this should be their clear and essential role. That is what is expected of them by involved organizations and individuals. Without it, I am mindful of Hamlet's tragic musings: "The native hue of resolution is sicklied o'er with the pale cast of thought."	We recognize that the public and decision makers would prefer a black or white answer as to risk, but reasonable people disagree on the EMF issue. Our policy anlysis contractors have demonstrated that decisions are nonetheless possible. Hamlet's dilema was not about uncertainty as to facts, but on the pros and cons of what to do about them.
Anders Ahlbom	Yes	We acknowledge your support for our point of view
David Bates	Yes; the document does an excellent job of presenting the uncertainties fairly. In fact, I think it is the best risk assessment analysis of a low risk outcome that I have encountered.	We acknowledge your support for our point of view
Carl Blackman	Has an adequate job been done presenting arguments for and against causality? I think the argument can be made stronger for causality if the NIEHS document had not taken as the exclusive and thorough review of the literature up to 1998. Other exhaustive reviews, for example by US Environmental Protection Agency and by NCRP, demonstrate consistent in vitro biological effects across laboratories together with demonstrations of unusual dose and frequency responses. Inclusions of these results would have strengthened the decision to prevent the lack of animal results, obtained with constant intensities and frequencies, from reducing the degree of confidence in the epidemiology data, obtained with complex exposures in the "real" world.	We note this comment. We will revisit the relevant part and amend them, if warranted.
Bowman	This question is extremely broad and begs an extensive, point-by-point reply, which	We acknowledge your support for our point of view

Question 7. Have we done an adequate job in presenting the arguments for and against causality or are we assigning weak arguments to the "con" or the "pro" position?		
	is not possible at this time. Although I criticized a few of your arguments in my specific comments, the document in general uses reasonable arguments to make both pro and con arguments as persuasively as possible.	
Shan Cretin	While you did a reasonable job for many of balancing pro and con arguments, I sometimes felt that the pro arguments were presented more strongly than they deserved. For example, there are underlying assumptions that studies are independent, but in fact, the study designs, measures and choice of covariates are quite similar across studies so that unrecognized flaws may (probably do) propagate through the literature. In general, scientists tend to think they know more than they do from individual studies. Physical constants measured in early experiments tend to share biases so that the ultimate number agreed on for the constant is often found to be outside the confidence intervals of the early studies.	We take notice of this comment. We presented the pro and con arguments uncritically, as if they were presented by advocates of the two points of view. The critical evaluation is presented in the third column. In the revised draft, we have explicitly addressed the possibility that several studies using a similar design may share a common bias.
John Dawsey *et al.*	You have presented the arguments, but you fail to assign sufficient value to the 'con' arguments and give too much credit to the 'pro' arguments. The analysis also lacks scientific rigor and does not give sufficient weight to key aspects of the scientific literature.	Although we tried to make our reasoning as transparent as possible, we accept the fact that interpretation of the pro and con arguments retains a degree of subjectivity. We acknowledge that. although you disagree with our conclusions, you agree with our presentation of the arguments.
Robert Goble	While we think the basic arguments are presented quite fairly, indeed remarkably so, it might helpful to describe up front the basic conundrum: that there is a persistence of epidemiological findings but a persistent failure to find supporting evidence elsewhere.	This is a good suggestion. We will try to modify the document accordingly.
Ben Greenebaum	As noted above, the arguments are not always consistent between chapters, presumably reflecting the assignment of various parts to various people and perhaps less review by the three reviewers and by others of the preliminary work. (Did the three review each of the sets of statements? If so, was this done before or after they formed their individual conclusions or afterwards? Different answers to these two questions could indicate a tendency toward different outcomes. Concern about the apparently small number of people involved—the 3 reviewers—shows up here.)	Indeed, each of the reviewers prepared the first draft of part of the document. However, each first draft was reviewed by the other members of the team BEFORE proceeding to the evaluation.
Leeka Kheifets	As you recall, I am not in favour of laying down "pro" and "con" arguments as you proposed.	
Patrick Levallois	Yes. This was done as it has never been before in all kind of experts group working on this kind of issue	We appreciate this comment
David McCormick	Specific comments are made in response to individual arguments.	

Question 7. Have we done an adequate job in presenting the arguments for and against causality or are we assigning weak arguments to the "con" or the "pro" position?		
Thomas McKone	The arguments for and against causality appear to be balanced. I did not get a sense that there was a systematic effort to assign weak or strong arguments to one or another side of the issue.	We appreciate this comment
Samuel Milham	An adequate job.	We appreciate this comment
Hal Morgenstern	I think one of the shortcomings of this report is that the pro and con arguments for evaluating the epidemiologic evidence are not thoroughly addressed. Refer to my comments above (i.e., c-n).	
Herbert Needleman	I thought the presentation of the arguments was strong. (see above) I have not encountered this format before.	We appreciate this comment
David Ozonoff	I think the presentation is outstanding in most respects. It is less clear only in the area which describes how the actual final degree of confidence was arrived at. I think some more words might be useful here. Description of the “pros” and “cons” is unusually thorough, thoughtful and sophisticated.	We appreciate this comment
Jack Sahl	Your pro arguments are speculative while the con arguments are based on observation. You consistently turn ‘ good ideas’ into evidence to support causality.	We disagree. We included speculative arguments in both the “pro” and “con” columns (in fact, in the review of a previous version, we have been accused by Dr Brown to include “strawmen” arguments in the “con” column. Our intent was to be as inclusive as possible and to evaluate the value of each argument on its merits.
Charles Quesenberry	Generally, the quality of the pro and con arguments seems balanced. Occasionally, an argument seems somewhat weak, but this was not the overall impression of the presentation.	We acknowledge your support for our point of view. We have included even weak arguments for completeness.
Rick Saunders	I would have taken a more critical view of the quality of some of the biological studies. In particular, the life-time animal studies and those on transgenic animals under the EMF Rapid programme were of really impressive quality; some of the others, which appear to have been given equal weight, were of a somewhat lower quality.	Our view is that even studies of an “impressive” quality are non ideally suited to study EMF effects, if they follow a protocol originallly designed for chemical agents.
David Savitz	The length and detail of the arguments is far beyond anything done before, and may well be more than is needed or helpful. It is hard for even those with a real academic interest to work through all the details, and it is unclear how much of a target audience there is for this level of detail.	We tried to be as comprehensive as possible, in order to avoid even the impression of overlooking arguments that might support one point of view or the other.
Joachim Schüz	I like the style of your presentation of the arguments very much. However, sometimes I got the impression that some of the arguments rely too much on single studies, and sometimes not even the best studies (one example is (F2) in Table	As noted above, we do not believe ourselves that all the points, either “pro” or “con” were very well founded. However, we included them all, to avoid even the impression of overlooking

Question 7. Have we done an adequate job in presenting the arguments for and against causality or are we assigning weak arguments to the "con" or the "pro" position?		
	8.2.2: the impact of selection bias should be discussed in the light of large studies that materially contribute to the EMF-leukemia-association, particularly the Linet study; e.g., the case-specular method for the Savitz and London study is completely irrelevant for the interpretation of the meta-analysis finding at 0.4μT by Ahlbom – and the studies that were part of this meaningful analysis that were affected by selection bias all show selection in the same direction).	arguments that might support one point of view or the other.
Asher Sheppard	I have not evaluated the "pro" and "con" positions in depth nor have I read all with care, but my impression is that there were no evident biases. There were, however, consistent views evident in the "comment and summary" statements drawn from the contrast of "pro" v "con".	
Gilles Theriault	I am in agreement with the arguments proposed. In general, they are well balanced and the "comment and summary" column sheds a neat light on where a reasonable person would stand. But very likely, opponents of the epidemiologists' view will question the value of the positions taken. The document, by its approach and the background of the three reviewers will be seen as the epidemiologists' "Bible" on health effect of EMF.	We appreciate this comment
Lorenzo Tomatis	I believe you have done a good job in presenting arguments in favor and against, although a certain degree of preconceived conviction may transpire. I would add that this is almost unavoidable, and welcome, for anybody who cares about public health.	We tried to be as objective and comprehensive as possible.
Jim Tucker	I believe you have done an adequate job of presenting the arguments for and against causality.	We appreciate this comment
Nancy Wertheimer	Based on the accumulated epidemiologic evidence, I think your assessments are well-balanced. Those strongly oriented against an effect by our lack of physical understanding may see you as too pro. I don't.	We appreciate this comment

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
Theodor Abelin	The overlap at 98 and 2% do not bother me, but the one at 50% seems awkward. A possibility would be to create a new category at the center of the scale, perhaps for the confidence range of 45-55%. But the disadvantage would probably be that it would give undecided reviewers a chance to avoid a clear statement. As someone from a non-English speaking country I was wondering whether your use of vocabulary is easy to understand for people with little education. How would the following alternatives be? Current phrase — Suggested alternative Virtually certain — Almost certain Highly probable — Very likely Possible >50% — Quite possible Possible <51% — Possible Very improbable — Very unlikely Virtually certain that it is not causal — Almost certain that it is not causal	We have considered all the comments received. We have eventually decided to adopt the classification described in Appendix 2.
Anders Ahlbom	Terminology is very difficult, as you state in your question and your approach is very good, still I think that these terms are virtually impossible to give precise meanings.	
David Bates	The only correction I would suggest to this Table is to use the phrase "probable" in describing the risks with a confidence range of 50-90%	
Carl Blackman	The plain language risk evaluation guidelines are not user friendly; how can they be improved? I think the present blurring of barriers between different categories is called for because the data do not allow for clear distinctions. Readers need to know that and learn to deal with it.	

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
Bowman	The phrases "50-90% Possible" and "10-50% Possible" definitely need to be modified. They are awkward, and use "possible" to mean "probability". The global warming report by the International Panel on Climate Change (Reilly et al., 2001) did far better: >99% Virtually certain 90-99% Very likely 66-90% Likely 33-66% Medium likely 10-33% Unlikely, etc. For public health purposes, however, I would replace the IPCC's adjectives with the familiar IARC terms "Probable" and "Possible".	
Shan Cretin	One problem with this table is that I would not call 2% "virtually certain"—I'd save that for a 995 in a thousand (or 5 in a thousand) chance. Another is that I would have a neutral mid band around 50. Words mean different things to different people, however, which is why we use numbers! My words would be: 99.5% virtually certain; 99.5 to 98% highly probable, 90-98% probable; 60 to 90 more likely to be causal than not; 40 to 60 equivocal; 10 to 40 less likely to be causal than not; 2 to 10 improbable; 0.5 to 2 highly improbable, 0.5 virtually certain not causal. Another way to approach putting words to numbers is to reinterpret the numbers to everyday events: The risk of being in an auto accident if a 30 year old drives X miles. (putting 30 year old in to take care of the fact that 16 year old boys or 87 year old women might have a different experience). Or the probability that there will be a 6,0 or larger earthquake in California in the next N minutes (hours, days?)	

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
John Dawsey *et al.*	There is no scientific justification for these categories. These are not consistent with the text used to describe the assessments of any other independent expert panel. For example, based on the same epidemiological data, the National Institutes of Environmental Health Sciences (NIEHS) concluded that: The scientific evidence suggesting that ELF-EMF exposures pose any health risk is weak. The NIEHS concludes that ELF-EMF exposure cannot be recognized at this time as entirely safe because of weak evidence that exposure may pose a leukemia hazard. The National Toxicology Program routinely examines environmental exposures to determine the degree to which they constitute a human cancer risk and produces the 'Report on Carcinogens' listing agents that are 'known human carcinogens' or 'reasonably anticipated to be human carcinogens'. It is our opinion that based on evidence to date, ELF-EMF exposure would not be listed in the 'Report on Carcinogens' as an agent 'reasonably anticipated to be human carcinogens.' NIEHS Director's EMF-RAPID Report to Congress, June 1999 For an international perspective, the U.K. National Radiation Protection Broad (NRPB) has a standing committee, chaired by an eminent epidemiologist, Sir Richard Doll. This committee has concluded: In the absence of any unambiguous experimental evidence to suggest that exposure to these electromagnetic fields was likely to be carcinogenic, the Advisory Group concluded that the findings of the epidemiological studies that had been reviewed could be regarded only as sufficient to justify formulating hypotheses for testing by further investigation. They provided no firm evidence of a carcinogenic hazard to either children or adults from exposure to normal levels of power frequency electromagnetic fields. National Radiation Protection Board, United Kingdom, March 2001 Regarding question 8 (above), CDHS should use statements that are similar to those of NIEHS or the U.K. NRPB. The scientific validity and reliability of expressing your views of risk based on 'categories of percentage likelihood' have not been established. If CDHS must use this approach, the full confidence range from the lowest of the low to the highest of the high should be indicated. For example, rather than stating an exact "xxxx", use a range, like, "It is between '20% and 90% likely' that EMF's at home or at work could cause a very small increased	Mr. Dawsey also opposed this approach when we formulated our Risk Evaluation Guidelines. He favors dichotomizing agents into thos "reasonably anticipated to be carcinogens" and those that are not. But the criteria for "reasonable" or the confidence that corresponds to this cutpoint are obscure. When we look at the agents that fit into the "reasonably anticipated" category, we find that they are ones we would characterize as "very probable" or extremely probable to be carcinogens. The fact that "reasonable" excludes "quite probable" and "moderately probable" agents is not made obvious to the public.

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
Robert Goble	The point of verbal descriptions is to provide an alternative to the mathematical representation: to be properly informative it will necessarily serve as a different sort of descriptor, conveying different informational value. Concern about the 50-51% overlap is only appropriate if the evaluators believe that they can make such a distinction, or that they know the balance of probability that well. Perhaps a bit more discussion of qualitative phrasing and why the numbers are not precise and do not reflect the exact nature of the judgements made would help more than trying to rephrase the categories.	

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
Ben Greenebaum	ANY set of phrases is going to be a problem. Prominently displaying the table of confidence range vs. phrase, supplementing it with a graphical view (divided bar or line), etc. is the only way to reduce the problem. You won't eliminate it. Suggestion below may or may not be better for a non-technical reader: Emphasize that boundaries are arbitrary and that position along 10-50 and 51-90 ranges reflects differences in degree of likelihood: Level of Confidence: 0% 2% 10% 50% 90% 98% 100% ←-->←----→←----------→←----------→←--------→←--> | | | | | 98-100%= | | | | | Virtually certain | | | | 90-98%=Highly likely | | | 51-90%= Possible, probably more than 50:50 chance | | 10-50%= Possible, probably less than 50:50 chance | 2-10% =Very unlikely 0-2%=Virtually certain that it is not a cause	We have adopted a scheme something like this.

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
Mark Israel	For the CDHS degree of confidence evaluation to be useful for scientists and the public, it needs to reflect an objective analysis of the weight of scientific evidence. To do so, the evaluation needs to be based on a clearly articulated set of scientific criteria for weighing the underlying evidence. This, however, is strikingly absent. The draft CDHS Report does not provide a description of any such criteria and, to the extent that any criteria were used, the draft Report does not explain how those criteria were applied to the scientific evidence. This failure to base the degree of confidence analysis on objective scientific criteria is a fundamental flaw in the draft CDHS Report. The use of so-called "plain language phrases" as labels for the various levels of "confidence" does not and cannot remedy this flaw. Without objective criteria for scientists and the public to evaluate, the degree of confidence analysis in the draft CDHS Report is not a scientific characterization of the relevant research.	We believe that similar criticisms could be levieed against the IARC process, in which membership in a category hingens on judgements on which adjective best describes a pattern of evidence, for example "limited" vs "inadequate".
Leeka Kheifets	I suggest changing "possible >50%" to "probable" and "possible <51%" to "not very probable."	
Patrick Levallois	50-90% = highly possible, 10-49 % slightly possible	
David McCormick	Suggested alternative terms are provided in the Table below. **See Table.**	
Thomas McKone	See Table below.	
Samuel Milham	There is a typo in the current phrase for the 90-98% confidence range. It should be highly probable, not highly probably. Any arbitrary ranking scheme will have problems. this one is as good as any.	
Hal Morgenstern	I would label the 50-90% confidence range as "probable" and the 10-90% range as "possible." Do not use probabilities (i.e., >50% and <50%) to further describe each confidence range, which is defined in terms of probabilities.	
Herbert Needleman	I have no difficulty with this classification.	

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
David Ozonoff	I think the division between "Virtually certain" and "Highly probable" is problematic. Isn't what you mean "I'm quite confident it's causal" and "I'm pretty certain it's causal" and not much more? Trying to assign numerical values here is pretty meaningless. Both of these categories are in the "above 90%" category, and trying to find a border in that range doesn't really mean much.	
Charles Quesenberry	The only way I can think of eliminating the overlap issue in the current phrasing of confidence is to either a) remove the ">50%" and "<51%" from the phrases, and reword to something like "Moderately probable", and "Moderately improbable"; or b) always include the range of certainty. I agree that the overlap with "Possible <51%" and "Possible > 50% is problematic. I suggest either using <=50% or < 50%.	
Rick Saunders	Perhaps for the confidence range 10-50% you could have 'less possible' and for 50-90% have 'more possible' rather than just 'possible' for both.	
David Savitz	I think that the phrasing and numbers are fine for clear communication purposes.	
Joachim Schüz	No comment. I felt that your phrasing was ok.	
Asher Sheppard	As noted above, I do not favor marrying numbers with seeming statistical precision to qualitative language when only qualitative terminology is justified. I recognize that the needs of the decision analysis drove an interest in producing numbers. The words used in the current phrases are in most cases just coded restatements of the numbers shown in the "confidence range" column. I offer as substitutes for the ranking of risks without an association with numerical values the following terms that, to the best of my ability, do not connote statistical concepts: *"Definite"; "Expected"; "Unexpected; "Remote";* and *"No Appreciable Risk"*. I would be just as happy to see only three values: *"Expected"; "Unexpected";* and *"Remote"*, because these should cover almost every realistic risk scenario. For example, a lifetime of heavy tobacco use is "expected" to cause a disease, but even adding up lung cancer, heart disease, pancreatic cancer, stroke, etc., it is not *"Definite"*. Some high asbestos exposures could be a counter-example, but I think the three terms *"Expected"; "Unexpected";* and *"Remote"*, capture most risks. One nice thing about having three categories is that it defies the split between greater than 50% or less than 50% that some segments of society put so much emphasis on.	

Question 8. Our Risk Evaluation guidelines (REGs) define some "plain language phrases" to express our degrees of confidence. However, when we actually applied them we found they were not problem free: a) Some of these phrases are not mutually exclusive. for example, Possible >50% overlaps "highly probable" and "virtually certain." In this case, the overlap is slight, but important, since it is about the "balance of probability." b) These phrases are grammatically awkward and they are not really "user friendly." How could we rephrase them, without violating the spirit of the REGs? please write any suggestions next to each phrase.		
Name of respondent	**Response to question**	**Staff reply**
Gilles Theriault	I am not troubled by your classification. A layperson will understand the language easily. The only change that I would propose is to replace Possible <51% by Possible <50% and Possible >50%, Possible ≥50%.	
Lorenzo Tomatis	I am not convinced of the usefulness of quantifying the confidence ranges in the way you propose. Although the IARC categorization in 1, 2A, 2B and 3 is far from perfect, I don't think that you reach much more by the percentages you propose. For instance, I don't see how one may decide when to keep out some of the 50-90% from the higher category? You just add one more category to those of IARC, and use numbers (percentages) instead of words. We have battled for years on those words, but nobody seems to be able to come out with a better classification, yet. The only reason in favor of using your percentages would be if they could actually help regulatory people to come to reasonable decisions. I am unable to evaluate that site of the issue.	It seems that the IARC category of "possible" includes things as varied as :"Coffee" that we would assign an 11% probable category and "fiberglass" that we would assign an 89% probable category. This is too broad a category
Jim Tucker	For this Table, I would be inclined to keep the words in column 2 as they are current written. However, I suggest that the number ranges in column 1 be changed slightly to avoid overlap, as follows: Confidence range — New confidence range >98% — (same) 90-98% — 90-98% 50-90% — 50-<90% 10-50% — 10-<50% 2-10% — 2-<10% <2% — (same)	
Nancy Wertheimer	I think your present wording is adequate.	

David McCormick

Confidence Range	**Current Phrase**	**Suggested Alternative**
>98%	Virtually certain	Virtually certain causality
90-98%	Highly probable [sic]	Highly probable
50-90%	Possible >50%	Probable (>50%)
10-50%	Possible <51%	Improbable (<50%)
2-10%	Very improbable	Highly Improbable
<2%	Virtually certain that it is not causal	Virtually certain lack of causality

Thomas McKone

Confidence range	Current Phrase	Suggested alternative
>98%	Virtually certain	*Virtually certain*
90-98%	Highly probably	*Highly likely*
50-90%	Possible >50%	*More likely than not*
10-50%	Possible <51%	*Possible*
2-10%	Very improbable	*Very unlikely*
<2%	Virtually certain that it is not causal	Virtually certain that it is not causal

APPENDICES

Appendix 1 - Qualitative Bayesian Evaluation

We coined the term "Qualitative Bayes Approach" to characterize a form of verbally justifying judgements about hazard that paid attention to the insights of Thomas Bayes, an 18th century mathematician. He pointed out that ones betting odds on a proposition after researching it was equal to the betting odds prior to having done the research multiplied by the likelihood of the research evidence if the proposition was true divided by the likelihood of the research evidence if the proposition was false (the relative likelihood). In our "Qualitative Bayes Approach" we elicited an expert judgement about a prior probability of hazard and after consideration of a carefully structured discussion of the evidence elicited an expert judgement on the posterior probability of a causal relationship. The carefully structured discussion should consider how much more (or less) likely the pattern of evidence would be if the risk hypothesis were true compared to the likelihood of that evidence if EMFs were safe. This consideration should be guided by a series of pre-agreed upon questions with graded (a lot less likely, somewhat less likely, about as likely etc) answers. The discussion should include pro, con and summary arguments. While this structured discussion should help organize the reviewer's judgement, after he/she weighs all the information in the usual way, it does not involve a mathematical combination of weights as would be the case in a quantitative Bayes evaluation.

The following table compares the "Qualitative Bayes evaluation to the traditional and to the Quantitative Bayes approaches to risk evaluation as to a number of characteristics.

Characteristic	Usual Method	Qual. Bayes	Quant. Bayes
Evaluates All Streams of Evidence	Sometimes	Yes	Focuses on Epidemiology, other streams influence prior
Elicits Prior Probability	No	Yes	Prior Dose Response Curve
Compares Likelihood of each element of the evidence under the hazard and non-hazard hypotheses?	No	Qualitatively	Quantitatively with many of the parameters subjectively elicited
Pro, Con and Summary Arguments to make rationale transparent?	No, most risk assessments are skimpy in justifying hazard categories assigned	Yes	Not unless a supplementary document were to accompany the model
Combines relative likelihoods mathematically to derive posterior?	No	No	Yes, but in some versions non-epidemiol. evidence is folded into the prior subjectively
Elicits an Expert posterior probability after considering all elements of the evidence?	No	Yes	No
Displays judgements of various judges separately	Usually strives for semblance of consensus	Yes	Technically Possible for different experts to elicit their own

			parameters
Frames median degrees of confidence as "Not a proven hazard"	Usually	No, reveals posterior probability	No, reveals posterior probability

Both the qualitative Bayes and the quantitative Bayes evaluations provide a posterior degree of confidence which, if in the range from 10% to 90% will not seem trivial to the general public and will stimulate policy discussions. The statements "there is no proven hazard" or "there is no consistence evidence" often used for this range of degrees of confidence, will not stimulate such discussions.

Thus, both the qualitative Bayes and quantitative Bayes methods pose risk communication "problems" for those who believe that society should not begin policy discussions until most scientists are more than 90% confident that a hazard exists. The traditional hazard identifications would pose the same "problem" if they routinely used more nuanced categories of hazard, that distinguished between, say 11% confident and 89% confident. As now framed they pose a risk communication "problem" for those who believe that policy discussions should begin before most scientists are extremely confident that a hazard exists.

Some of the commentors who disagreed with the posterior degrees of confidence of our three reviewers seemed to believe that the deviation from their own judgements was an inevitable result of the qualitative Bayes approach. Greenland in particular felt that we had double counted some concepts and had not given enough consideration to the possibility of various combinations of bias, confounding and chance. We will address his specific concerns separately.

We see no reason why the qualitative Bayes format of risk evaluation will always produce higher degrees of confidence of hazard than the traditional or quantitative Bayes format. The fact that our three reviewers, who had ample opportunity to argue with each other, still varied in their judgements is testimony for this point. We have no doubt that our critics could use the qualitative Bayes format to make their points. Some of the physicists would use very low priors; the toxicologists would make a case that the animal studies conveyed a relative likelihood less than 1.0 to pull down their confidence. Reviewer 2, interpreted the pro and con arguments in a way that is not in conflict with Greenland's points and came up with a posterior degree of confidence for childhood leukemia not very different from Greenland's' 50%.

In a contentious area such as EMFs, we doubt very much that any of these three styles of risk evaluation would force a consensus among subject matter experts who weigh and interpret the several streams of evidence differently. Even in the quantitative Bayes model experts will use different priors and will elicit different subjective relative likelihood parameters.

In the last draft of the risk evaluation we plan to alter the executive summary, chapters 1, 2 and 7 to make the above points clear to the reader.

Appendix 2 - Comparison between this and the IARC and NIEHS Evaluations

How different is this evaluation from the NIEHS and IARC findings? The following table compares the evaluations on the part of the NIEHS Working Group, the IARC and this Program.

Health Outcome	NIEHS	IARC	DHS
Childhood leukemia	2B	2B	2B to 1
Adult leukemia	2B	inadequate	2B to 1
Adult brain cancer	inadequate	inadequate	2B
Miscarriage	inadequate	NA	2B
ALS	inadequate	NA	2B
All other	inadequate	Inadequate or NA	inadequate

It is clear from the table that, when applying the IARC guidelines, the DHS reviewers agreed with IARC and NIEHS reviewers that in many cases (e.g., childhood brain cancer and male and female breast cancer) the evidence was inadequate to reach a conclusion. One of the DHS reviewers agreed with the IARC and NIEHS on childhood leukemia. Two of the reviewers agree with NIEHS, but not with IARC, on adult leukemia. Otherwise, the DHS reviewers regard the EMFs association more likely to be causal than either IARC or NIEHS did.

There is a wide range of opinions in the scientific community as to the probability that EMFs cause health problems. The DHS reviewers provided numerical values for their degrees of confidence that risk of various diseases could be increased to some degree by EMF exposure. Other researchers have rarely packaged their judgements in this way, but judging by one such exercise that we conducted (Bioelectromagnetics 2001) reasonable scientists can have different ways of interpreting the data with resulting different degrees of confidence. We assume that many members of the NIEHS and IARC working groups, if forced to provide a numerical degree of confidence would have given lower numbers that the CDHS group did.

Part of the reason for the discrepancies in the DHS reviewer's IARC classification choices can be traced to differences in the two evaluation processes. Reviewer 1 participated in both the DHS and IARC review and found it impossible to express the same opinion in the two processes, although his view had not changed in the meantime. This was due to the two-tier voting used in the IARC process and on the aim for consensus. The members of the working group were asked to vote separately on animal and human evidence. In both cases, reviewer 1 voted against the majority. Although a sizable part of the working Group believed that there was limited animal evidence indicating a possible cancer risk, their opinion was not carried past that point of the process. Since the majority regarded the animal evidence as "inadequate", when the final vote on the overall evaluation was taken, the option posed to the working group's members were the majority positions, that is , that animal evidence was inadequate and epidemiological evidence was limited. According to the guidelines, these two majority positions resulted automatically in a Group 2B classification. Therefore a class 2A or class 1 were not even considered as options to vote on, even if individual reviewers, such as Reviewer 1 might have so voted. In the DHS process each reviewer voted as if his or her position was a consensus- of- one position. The goal was to make any differences transparent to decision makers.

With regard to adult leukemia, the IARC's evaluation differs from the NIEHS and the California evaluation because of the way epidemiological evidence was considered. Almost all the evidence on adult leukemia comes from occupational studies. The Epidemiology subgroup at the IARC meeting regarded most of these studies as of poor quality, with within- and between-study inconsistencies. Most of the evaluation centered on the most recent large studies (Sahl, Savitz and Theriault), which contradicted each other.

Our evaluation considered the whole body of studies, residential and occupational (which our "sign test" approach allows us to do). While we acknowledge that many of the studies have serious limitations, neither we, nor other reviewers, have

identified fatal flaws. For example, there is no evidence to suggest that the use of crude exposure assessment surrogates, while virtually certain to influence the quantitative estimate of risk and to frustrate any attempt to explore the dose-response relationship, introduced an upward bias in the reported association. On the contrary, the limitations of the studies may well be responsible for the inconsistencies between them. And while these inconsistencies do exist, they are not as common as the IARC evaluation may suggest. The Kheifets meta-analysis concludes that the body of epidemiological evidence shows a slight but statistically significant increase in risk. From a binary outcome stand point, the studies with a relative risk estimate >1 are more than twice as numerous as those with a RR<= 1.

Another possible reason for the discrepancy between the DHS and the IARC and NIEHS evaluation is possibly due to the way the membership of the working groups were defined. In our case, the reviewers were defined a priori as the core scientific staff of the Program (although independent input was sought).

In the other cases, the working groups were constituted ad hoc. As noted elsewhere, since the EMF research community is small and views of its members are fairly well known, it is hard to select a small group without influencing the outcome. As a result, organizers try to include members from each side of the debate. This is an excellent solution, if the evaluation is carried out through extensive debates and discussions, conducted in good faith, with a shared desire to come to a consensus opinion, irrespective of its potential social and economic consequences. This was the original approach used by IARC (Tomatis, private communication). However, in time, the pressure to conclude the evaluation within a short period of time and the difficulty of separating scientific objectivity from external pressures, led to abandoning the discussion format in favor of the voting system. At the meeting to draft he EMF monograph (June 2001), the vast majority of the plenary session time was dedicated to reviewing the draft chapters prepared ahead of time but designated committee members with maybe 10% of the time allowed for discussion. Whenever a paragraph precipitated a controversial discussion , a common way out was to propose the deletion of the offending paragraph, a proposal that the time pressured working group members were usually glad to adopt.

Given the deliberate range of opinions in the working group, voting invariably leads to a "middle of the road" opinion.

In contrast to this process, the DHS reviewers spent innumerable hours and days, over a period of years and in consultation with independent consultants, to try to resolve their differences. Even so, and although two of the DHS reviewers were bureaucratically subordinate to the third, substantial differences remained and are openly revealed in this evaluation.

Where there is disagreement between this and other recent evaluations, it is difficult to gauge how wide the gap is. According to Dr Rice, Chief of IARC's Carcinogen Identification and Evaluation Unit, "If IARC were to say that an exposure is in Group 2A, probably carcinogenic to humans, that would mean that the evidence is just a little short of certainty that the exposure in question has actually caused human cancer. . . Group 2B is the lowest level of identifiable carcinogenic hazard in the IARC system.". Although Dr Rice stresses that the IARC classification is not based on a probabilistic model, it would not be unreasonable to compare a 2A classification to DHS's "very probable (>98%)" or "extremely probable (90 to 98%)", while the 2B group appears to encompass most of the remaining DHS's categories "quite probable", "moderately probable" and "quite improbable", spanning probabilities from 10 to 90% possible, excluding only the DHS categories "very probable that they are safe" and "extremely probable that they are safe" (probably corresponding to IARC's Group 4). We need to stress that this comparison is entirely our speculation and would not be necessarily accepted by IARC. However, the fact that the 2B classification includes coffee (which most people regard as safe) and fiberglass (which many people regard as a very likely carcinogen) supports our conclusion that the IARC 2B group spans a very wide range of probabilities and is to be regarded as a "limbo" in which agents are placed, when they their hazard cannot be confirmed nor ruled out.

Another fact that needs to be stressed is that the IARC Group 3 classification "inadequate evidence" is not equivalent to claiming that evidence does not support causal association or possible risk (Kavet's comments, Table 2) or that "there is no reason for concern" as stated in the NIEHS report to Congress [GET EXACT QUOTE], (but not in the NIEHS Working Group Report). For example, the IARC Working Group did not dismiss or explain away the substantial body of evidence of an association between EMF and adult leukemia and did not reject Kheifets' meta-analytical result. It simply considered that more evidence was required to clarify the current inconsistencies. It may not take much more evidence of an association to change the Group 3 classification to a Group 2B.

Our willingness to express a numerical opinion when others had chosen not to do so derives from the fact that our program supported two decision analytic policy analyses that required us to express our judgements as degrees of confidence. The fact that our IARC classifications were sometimes (but not always) higher than IARCs may partly be due to our use of the sign test (discussed elsewhere) which does not require obtaining raw data (as for a pooled analysis) and allowed us to

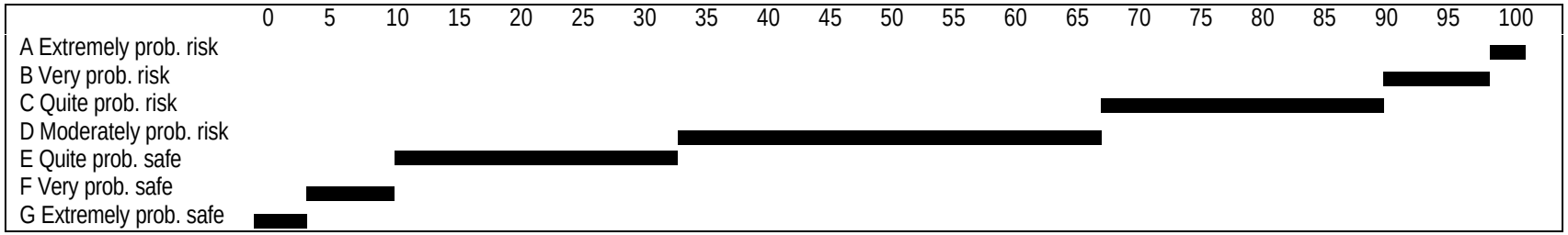
include in the pool of data to be evaluated results that are too heterogeneous for a meta-analysis. Therefore, in most cases the pool of data we examined and paid attention to was more comprehensive than those used by other evaluators.

Another reason is that we considered the existence of some experimental mechanistic results at low intensity magnetic fields as weakening the argument that the energy of environmental fields is too low to have any significant biological effects.

Sometimes evaluations are part and parcel of a policy recommendation (they may even include regulatory recommendations in the conclusion. For example the NIEHS report to congress concluded with a recommendation that the utility industry continue (sic) to use low cost ways to avoid producing high magnetic fields). Many evaluations aim to avoid false positive pronouncements, as seems to be the case with IARC judging by comments from Dr Jerry Rice: "….the IARC Monographs system of carcinogenic hazard evaluations is deliberately a very conservative one. There are many carcinogenic hazards in the human environment that are very real indeed, and control of exposures to those hazards is extremely important for public health. To accomplish this, it is necessary that carcinogenic hazards be correctly identified. We must avoid mis-directing public attention to any exposure of any kind that may be perceived as a hazard, but in fact is a misplaced concern. (Dr Jerry Rice, letter to Vincent DelPizzo, Aug 10 2001).. The California Department of Health Services three reviewers have packaged their differing degrees of confidence about causality in a way that can be used in the decision analytic models prepared for the program. It has pointed out that the policy implications of this range of confidences depends on the policy framework of the decision-maker: libertarian, utilitarian, virtual certainty or social justice. The public regulatory process will determine which one or which mixture of these frameworks will apply to govern policy. Thus the CDHS risk evaluation is packaged to facilitate decision making but separates risk assessment from risk management.

Our evaluation would not serve its purpose if the reviewers moderated it to avoid being perceived as alarmist, or decided to err on the side of safety. If we said, for example, that we judged that there is a 40% probability of EMF being harmful, we believe that the evidence suggests that they are probably safe. Therefore, we would be in agreement with other reviewers who might say that there is "no convincing evidence" that EMFs are hazardous. However, in our opinion that latter expression implies a default position that EMFs are safe until proven hazardous. We do not have such a default position: when we express a judgement on the probability of EMFs being hazardous we also express a belief that there is a complementary probability that EMFs are safe. This was implied in Draft 3, but will be made explicit in Draft 4 (see Table below)

Qualitative evaluation *Are EMFs at home or at work safe?* *or do EMFs increase the risk of..... to some degree?*	**Equivalent quantitative estimates** *Probability of EMFs at home or at work increasing risk to some degree*	*Approximate midpoint odds of EMF at home or at works increasing risk to some degree*
A) Extremely probable that research will show that they increase the risk to some degree (that is, extremely improbable that they are safe)	>98%	99 to 1
B) Very probable that research will show that they increase the risk to some degree (that is, very improbable that they are safe)	90 to 98%	16 to 1
C) Quite probable that research will show that they increase the risk to some degree (that is, quite improbable that they are safe)	66 to 98 %	5 to 1
D) Moderately probable that research will show that they increase the risk to some degree (but also moderately probable that they are safe)	33 to 66%	1 to 1 (2 to 1- 1 to 2)
E) Quite probable that research will show that they are safe (that is, quite improbable that they increase the risk to any degree)	10 to 33%	1 to 5
F) Very probable that research will show that they are safe (that is. very improbable that they increase the risk to any degree)	2 to 10%	1 to 16
G) Extremely probable that research will show they are safe, (that is, extremely improbable that they increase the risk to any degree)	< 2%	1 to 99



When viewed in this light, we believe that the discrepancies between our and other evaluations are less puzzling than they appear at first sight.

Appendix 3 - The sign test

The sign test (also called "vote counting") is not a new statistical method. It is less efficient and informative than a full-fledged meta-analysis or pooled analysis, and therefore it is rarely used. However, we believe that in the case of EMF, it also offers advantages.

With EMF, we are still trying to determine whether there is a risk resulting from exposure. Obviously, if this proves to be the case, it is desirable to estimate the size of the risk and the dose-response functions, information that can only be obtained by meta- or pooled analyses. However, we believe that to get past the first step, the sign test is very useful. Note that none of the existing meta or pooled analyses available include all the studies in the literature. In some cases (Greenland 2000), some data were not made available. In other cases, studies are so different (e.g., occupational and residential studies) that they cannot be combined in a meaningful way. In other cases doubts exist on the quality of some studies, which where therefore excluded from consideration (note that to exclude a study because of identified fatal flaws is essential, but to exclude a study because of suspected limitations may result in the exclusion of valid data).

The sign test is a simple statistical test that determines whether a pattern of binary results is too extreme to be consistent with the null hypothesis. By way of illustration, we discuss below the application of this method to the childhood leukemia studies:

We have considered 18 studies of childhood leukemia (the most recent one by Shuz et al is consistent with the pattern, but did not meet our deadline for inclusion). We took care to include only studies that had no identified fatal flaws, to include each study once and only once, even if the paper contained several risk estimates, for example for measured fields and for wire coding. For studies with measured or calculated fields, we only used one cutpoint, chosen a priori as the same cutpoint used by Greenland et al (2000), i.e., 3 mG. For non-measurement studies, we used a simple high vs. low classification and we expressed the result as a simple dichotomy (RR > 1 or RR #1).

Study #	Author	Country	Risk estimate	Binary outcome for >0.3 μT
1	Coghill	UK	no controls	?
2	Dockerty	New Zealand	no controls	?
3	Feychting	Sweden	4.44	+
4	Linet	USA	1.51	+
5	London	USA	1.53	+
6	McBride	Canada	1.42	+
7	Michaelis	Germany	2.48	+
8	Olsen	Denmark	2.00	+
9	Savitz	USA	3.87	+
10	Tomenius	Sweden	1.41	+
11	Tynes	Norway	no cases	?
12	Verkasalo	Finland	2.00	+
13	Green	Canada	1.23	+
14	UK	UK	0.97	–

			Non-measurement studies	Risk for the high exposure group
15	Wertheimer	USA	1	+
16	Fajardo	Mexico	>1	+
17	Coleman	UK	>1	+
18	Petridou	Greece	>1	+

For the sign test we only need to consider but only at the binary result in the last column, where + marks a risk > 1 and - a risk ≤1. Of the studies considered, 3 had no subjects in one of the exposed groups, so that no risk estimate could be calculated. Therefore, we are left with 15 trials and only 1 "success" (i.e. a point risk estimate ≤1). If there is no true underlying association there is only a 0.03% probability of observing a pattern of results as, or more extreme than that.

Kavet notes (correctly) that the sign test "does not reflect the inherent strengths and limitations of epidemiologic data". Is this a problem? We reviewed each study individually, before accepting them for the sign test. All of the studies considered were included in either the Greenland or Ahlbom pooled analyses or both.

The sign test does not consider the size of the individual studies. This is indeed a limitation. Does the test give small studies undue weight? No. All we need to properly apply the test is a series of trial with binary outcome. Even if the trial consisted in comparing the exposure of one case and one comparable control (with the outcome being determined by who has the highest exposure), this would be perfectly acceptable. Does the sign test give due regard to large studies? No. A trial including hundreds or thousands of subjects, has no more weight than the trial consisting of one case and one control. In this sense, as correctly pointed out by Greenland, the test is inefficient, because it does not use all the available information. This is no big problem if the large studies are "on the winning side", i.e., consistent with the smaller studies. However, suppose one has a large study with a null result and many studies with "positive" (i.e., RR > 1) results, the sign test would conclude that there is a real association. A meta-analysis would give more weight to the large study and conclude that there is no association. There are only two possible explanations:

1. The large study could be biased towards the null. Unless the analyst can spot the bias, the meta-analysis's result would be similarly biased, specifically because the biased study is the largest one and carries more weight. This does not happen with the sign test.

2. The suggested scenario is unrealistic. If there is no true association, why should the many small studies consistently be positive? Small studies are intrinsically unstable. Therefore, if there is no true association, in reality they cannot be expected to be consistently positive, unless one assumes that all the small studies are consistently biased upward. We reject this suggestion for two reasons: a) it "assumes facts not in evidence", to use legal jargon; and b) bias is not correlated to study size, therefore we have no reason to believe the large study to be more likely to be bias-free than the many small studies.

Finally, one could object to the fact that the sign test does not consider the strength of the association. Suppose one has a large number of studies, consistently showing an extremely small risk (say, RR=1.05). According to the sign test, one should be very confident that a true association exists, while a meta-analysis or a pooled analysis, which are more efficient statistical tools, would probably still yield a risk estimate with a confidence interval including one. Accordingly, one could argue that this example shows that the sign test is misleading. We argue that this is an example of a case in which the sign test is superior to a meta analysis. If the association is very weak, a pooled analysis requires very large sample sizes to produce sufficiently tight confidence intervals. A sign test, only needs to determine whether the pattern of results is too extreme to be due to chance or not. There is no reason to expect a series of studies (good or not so good, large or small) to yield results which are consistently above one, unless one of the two following alternatives exist:

- There is an association (causal or not)
- The number of the studies is too small to rule out chance as the explanation, in which case, the sign test would return an appropriately large p-value to warn us of this possibility

Appendix 4 - Possible association between SES and non-participation in two San Francisco Bay Area studies

The nested case-control study of EMF personal exposure and miscarriage by Lee et al. (2002) does provide some weak evidence that non-participation bias may be responsible for the association between wire code and the risk of miscarriage. In that study, a number of subject refused to have their home and personal measurement taken. To evaluate the effect of this refusal they wire coded virtually all cooperating and non cooperating cases and a sample of all controls. The wire code association among the participating subjects could then be compared to the inferred association in the entire cohort. From Table 5 in Lee et al. (2002) we have:

Wire Code	Participant Case	Participant Control		Non- Part. Case	Sample of Controls
VHCC	26 (15.2%)	68(12.5%)		9	48(12.1%)
Other	145(84.8%)	474(87.5%)		115	348(87.9%)

OR = 1.25 (0.74-2.09)

Our best estimate of the full cohort association with wire code can be derived from the full case series (participant and non participant) and the random sample from the drive- by of all controls.

Wire Code	All Cases	Sample of Non Cases
VHCC	26+9 = 35 (11.9%)	48
Other	145+115 =260 (88.1%)	348

OR = 0.976(0.60-1.59)

So a selection bias produced a 1.25/0.976 =1.28 fold increase in the apparent risk conveyed by VHCC wire code. However, these two risk estimates are well within each others confidence intervals. Judging by the column percentages non-participation bias, if present, was due to VHCC cases being relatively more likely to participate, rather than VHCC controls being less likely to participate, as suggested by Hatch and Bracken. Controls seem to have participated regardless of VHCC status.

At best, these data suggest that selection bias like this might explain the persistent association between wire code and childhood leukemia, but not the association between measured fields and leukemia because nearby power lines are not the only source of magnetic field exposure. Furthermore, this modest non-participation bias (if any) with regard to wire code was not related to SES. When we compare the parent cohort to the nested case-control study, we find no evidence that the distribution of incomes among subjects in the nested case-control study (which required agreement to participate) is different than that among the whole cohort from which these subject were drawn (Lee, Neutra, Hristova et al. personal communication 2001)

Prospective Parent (Interview) Cohort

Income (thousands)	Case	Control	OR
<$20	42	408	1.0
$20-$40	138	1405	0.95
$40-$60	165	1484	1.08
$60-$75	71	541	1.27
>$75	74	693	1.04

Chi Square for Linear Trend 1.44 p= 0.23

Nested Case Control Study

Income (thousands)	Case	Control	OR
<$20	6	20	1.0
$20-$40	33	122	0.90
$40-$60	50	63	1.07
$60-$75	31	71	1.46
>$75	42	133	1.05

Chi Square for Linear Trend 0.82 p =0.36

Furthermore, we evaluated the association between income and wire code for the 541 controls and 174 cases in the nested case control study for whom we had complete information on both variables. We compared those with VHCC wire codes to those in other categories as to the same 5 category income scale presented elsewhere. The odds of having a VHCC home did not increase progressively as income fell either for cases or controls. From highest to lowest income, the relative odds for cases was 1.00, 0.9, 1.9; 1.1, 0.0; Chi Square for trend= 0.05, p= 0.8. The analogous set of relative odds for controls was 1.00, 1.1,1.1,1.3,0.8, Chi Square for linear trend was 0.2, p value = 0.7. Thus the slight non-participation bias on wire code in the nested case control study was not reflecting an effect of income, since income was not related to wire code in this study.

Recent analyses of the childhood leukemia data (Greenland 2000, Ahlbom 2001) have dispelled the notion the wire codes were the stronger predictor of risk, nor was wirecode a risk for miscarriage in the Li 2001 and Lee 2002 studies. Measure or calculated fields turned out to be more consistently associated with disease. The question then is: Is income a predictor of personal magnetic field exposure?

The best data we have on this question comes from a prospective cohort study of personal magnetic field and subsequent miscarriage in urban San Francisco (Li 2001). Subjects were recruited early in pregnancy prior to miscarriage, so selection bias could not have occurred.

In the Li et al prospective study the correlations between income and time weighted average magnetic fields (TWA) and maximum magnetic field exposure(MAX) were

Variable	Pearson Correlation	p value
TWA	0.034	0.30
Maximum exposure	0.038	0.25

The large number of subjects (nearly 1000 pregnant women) allowed Li (personal communication) to display the 25th, 50th and 75th percentile of exposure within quintiles of income. There was no consistent trend with income.

In the large nested case control of measured magnetic fields and miscarriage by Lee et al. (2002) one can examine the correlations between personal exposure and family income in a 15 point scale for the 663 controls. Once again there is little correlation:

Variable	Pearson Correlation	p value
TWA	0.025	0.53
Maximum exposure	0.04	0.31

Thus, in the only two studies which allow one to assess the association between personal exposure and family income, there seems to be little or no association.